Computers & Operations Research 43 (2014) 90–99

Contents lists available at ScienceDirect

Computers & Operations Research

journal homepage: www.elsevier.com/locate/caor

Scenario grouping in a progressive hedging-based meta-heuristic for stochastic network design

Teodor Gabriel Crainic [a,b,*], Mike Hewitt [c], Walter Rei [a,b]

[a] CIRRELT, Canada
[b] Département de management et technologie, École des sciences de la gestion, U.Q.A.M., Canada
[c] Department of Information Systems & Operations Management, Quinlan School of Business, Loyola University Chicago, United States

ARTICLE INFO

Available online 29 August 2013

Keywords:
Stochastic programs
Network design
Progressive hedging
Scenario clustering
Machine learning

ABSTRACT

We propose a methodological approach to build strategies for grouping scenarios as defined by the type of scenario decomposition, type of grouping, and the measures specifying scenario similarity. We evaluate these strategies in the context of stochastic network design by analyzing the behavior and performance of a new progressive hedging-based meta-heuristic for stochastic network design that solves subproblems comprising multiple scenarios. We compare the proposed strategies not only among themselves, but also against the strategy of grouping scenarios randomly and the lower bound provided by a state-of-the-art MIP solver. The results show that, by solving multi-scenario subproblems generated by the strategies we propose, the meta-heuristic produces better results in terms of solution quality and computing efficiency than when either single-scenario subproblems or multiple-scenario subproblems that are generated by picking scenarios at random are solved. The results also show that, considering all the strategies tested, the covering strategy with respect to commodity demands leads to the highest quality solutions and the quickest convergence.

1. Introduction

Network design models define an important class of combinatorial optimization problems with a wide gamut of applications. These problems naturally appear in various forms in the planning of complex systems, e.g., logistics, transportation and telecommunications, at the strategic, tactical and operational levels. In such contexts, network design models are used to produce plans that define the structure, services, allocation of resources, or adjustments to be applied to the respective networks. Such plans are then used for varying periods of time depending on the decision level considered. In the case of either strategic or tactical planning, decisions are made for relatively long periods of time. Therefore, managers responsible for such planning decisions generally face uncertainty (i.e., stochastic parameters) at the moment when plans are being drawn.

Demand usually entails a certain level of uncertainty for network design problems, defining stochastic parameters relative to, e.g., origins, destinations, or demand volumes. In this paper, we assume that origins and destinations are known, but volumes are uncertain. To account for such uncertainty, forecasting is traditionally used to obtain estimates in replacement of stochastic parameters. Managers may also apply some form of sensitivity analysis to provide alternative plans and networks. This type of approach can lead to arbitrarily bad solutions, however [31]. Moreover, recent studies have shown that cost-effective networks obtained in stochastic settings are structurally different than the ones obtained in deterministic settings [19,28].

Stochastic programming has become the methodology of choice to properly account for uncertainty in planning problems. The goal of stochastic programming approaches for network design is to build a single design that remains cost-effective when different demand realizations are encountered. To do so, uncertainty in demand is typically modeled with a finite set of scenarios, which must be generated with care to ensure that they, collectively, closely approximate the uncertainty in the planning setting (e.g., [8]). Once the appropriate set of scenarios is generated, a two-stage stochastic network design problem is generally solved, the first stage modeling the choice of design (e.g., selection of transportation services and schedules to operate during the next season or of components to include in the logistics network), the second modeling its cost-effectiveness in servicing the demand (e.g., routing) for each scenario. See [18] for a thorough review of such models.

Even though such problems have recently received increased attention from the scientific community, they remain notoriously hard to solve. The reasons for this are that, on one hand, deterministic network design problems are NP-Hard in all but trivial cases and, thus, formulations of even moderate size are generally difficult

* Corresponding author at: Département de management et technologie, École des sciences de la gestion, U.Q.A.M., Canada. Tel.: +1 514 343 7143; fax: +1 514 343 7121.
E-mail addresses: TeodorGabriel.Crainic@cirrelt.ca (T.G. Crainic), mrheie@rit.edu (M. Hewitt), Walter.Rei@cirrelt.ca (W. Rei).

0305-0548/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.cor.2013.08.020

to address and, on the other hand, modeling uncertainty with scenarios generally yields very large instances. Thus, much like with the deterministic case, we need heuristic methods for producing high-quality designs for realistically sized instances, and we need efficient methodology to address problem settings involving large numbers of scenarios.

Strategies decomposing the set of scenarios have been used to address these challenges and build solution methods for stochastic combinatorial optimization problems. Meta-heuristics based on this approach and the progressive hedging idea [23] have proved computationally successful in several such settings, including stochastic network design (e.g., [8,13,20]). Yet, the instances addressed are still of relatively modest dimensions, as the scenario decomposition used by these meta-heuristics generated single-scenario subproblems, each a NP-Hard deterministic network design formulation.

Revisiting the scenario-decomposition strategy by grouping scenarios to yield multi-scenario subproblems thus appears methodologically interesting as it could reduce significantly the number of subproblems, at the price of an increase in the difficulty of addressing each subproblem. The interest is also motivated by recent studies showing that exact methods using scenario decomposition to address integer stochastic problems can be significantly improved by applying a decomposition strategy based on scenario groups [10,11]. The general idea was to adapt the Lagrangian relaxation to produce subproblems defined on randomly defined groups of scenarios. Escudero et al. [11] observed improved lower bounds obtained by solving the associated Lagrange dual problem, and have numerically shown that the computational burden of solving the Lagrange dual problem is also reduced when grouping is used. Although grouping scenarios has clearly proven to be efficient in this case, no studies that we are aware of have focused on how such strategies should be defined and scenario groups be created. Furthermore, no scenario-decomposition strategy based on scenario grouping has yet been applied in the context of meta-heuristics.

The goal of this paper therefore is twofold. First, to propose a systematic methodological approach to build strategies for grouping scenarios as defined by the type of scenario decomposition (partition or cover), type of grouping (similar or dissimilar scenarios), and the measures specifying scenario similarity. Second, to study these strategies experimentally in the context of stochastic network design in order to characterize their behavior and recommend how to use the proposed methodology. To perform this study, we introduce a new progressive hedging-based meta-heuristic for stochastic network design that solves subproblems comprising multiple scenarios as defined by the proposed scenario-grouping methodology. Notice that, while the scenario-grouping strategies we propose are derived in the context of addressing stochastic network design problems, they can be applied to other stochastic programs as well.

We evaluate the performance of the proposed strategies for grouping scenarios through the effectiveness of the resulting progressive hedging meta-heuristic. We use a state-of-the-art commercial solver to solve subproblems to optimality. This enables us to focus the analysis on the benefits of solving multi-scenario subproblems and reduce any potential noise that may occur by solving them with a heuristic method. The results show that solving multi-scenario subproblems based on a random partition of scenarios often enables the progressive hedging meta-heuristic to achieve a solution that is 25% better than when single-scenario subproblems are solved, and in fewer than half the iterations. Moreover, the results also show that, compared to grouping scenarios randomly, partitioning scenarios with the grouping strategy we propose can enable the meta-heuristic to obtain a solution that is 16% better. The number of iterations and time required to achieve this result is always reduced, often by half. Finally, the results show that a covering of scenarios enables the meta-heuristic to find solutions that are 16% better than partitioning them. Thus, covering produces an overall improvement in solution quality of approximately 27% with respect to the original strategy of solving single-scenario subproblems.

The rest of this paper is organized as follows. We recall the two-stage formulation of the stochastic network design problem in Section 2 and review the related literature. We analyze the issue of grouping scenarios in Section 3 and introduce the various strategies we propose. We start Section 4 dedicated to the description of the computational experiments we performed by introducing the new progressive hedging-based meta-heuristic that can solve subproblems comprising multiple scenarios. We then proceed to computationally study the behavior and performance of the proposed methodology for grouping scenarios. We draw conclusions based on these experiments and outline future efforts in Section 5.

2. A brief tour of stochastic network design

We first recall the two-stage formulation of the stochastic network design problem and review the main applications and models present in the literature (Section 2.1), and then provide a general description of the algorithmic strategies used to develop the solution methods proposed for these problems (Section 2.2).

2.1. Stochastic network design models

Network design models entail two general groups of decisions: design decisions, that define the structure and characteristics of the network, and flow decisions, which relate to how the network is used to perform the operational activities considered, see [9]. When using the *a priori* approach [6] in a stochastic setting, decisions are made in stages according to when stochastic parameters become known. Problems are therefore formulated by defining which decisions are taken before all information is available (first stage decisions) and which decisions are made afterwards (second stage decisions and onward). Traditionally, in the case of stochastic two-stage network design models, design decisions define the first stage (i.e., the *a priori* solution) and flow decisions the second (i.e., the available recourse), see [18].

Formally, given a directed network with node set N, arc set A, commodity set K, and scenario set S, we wish to

$$\min \sum_{(i,j) \in A} f_{ij} y_{ij} + \sum_{s \in S} p_s \left(\sum_{k \in K} \sum_{(i,j) \in A} c_{ij}^k x_{ij}^{ks} \right)$$

subject to

$$\sum_{j \in N^+(i)} x_{ij}^{ks} - \sum_{j \in N^-(i)} x_{ji}^{ks} = d_i^{ks} \quad \forall i \in N, k \in K, s \in S, \tag{1}$$

$$\sum_{k \in K} x_{ij}^{ks} \le u_{ij} y_{ij} \quad \forall (i,j) \in A, s \in S, \tag{2}$$

$$y_{ij} \in \{0,1\} \quad \forall (i,j) \in A, \tag{3}$$

$$x_{ij}^{ks} \ge 0 \quad \forall (i,j) \in A, k \in K, s \in S, \tag{4}$$

where y_{ij} indicates whether arc (i,j) is installed in the network, f_{ij} is the cost (often called the fixed charge) of doing so, ${x_{ij}}^{ks}$ is the amount of commodity k's demand that flows on arc (i,j) in the resulting solution for scenario s, and ${c_{ij}}^k$ is the cost per unit of demand flowed on arc (i,j). Constraints (1) ensure that in each scenario s, each commodity's demand may be routed from its origin node to its destination node. Constraints (2) ensure that the same design is used in each scenario, and that arc capacity (u_{ij}) is never violated. When $d^{ks} < u_{ij}$, the disaggregate inequalities $x_{ij}^{ks} \le d^{ks} y_{ij}$ can be added to the formulation to strengthen its linear relaxation. We refer to this problem as the $CMND(S)$; its optimal solution is a single design that is cost-effective for all scenarios.