Operations Research Letters 44 (2016) 191–195

Contents lists available at ScienceDirect

Operations Research Letters

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

A note on hierarchical hubbing for a generalization of the VPN problem

Neil Olver *

Department of Econometrics & Operations Research, Vrije Universiteit Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, Netherlands
Centrum Wiskunde & Informatica (CWI), Science Park 123, 1098 XG Amsterdam, Netherlands

ARTICLE INFO

Article history:
Received 9 June 2015
Received in revised form
31 December 2015
Accepted 31 December 2015
Available online 8 January 2016

Keywords:
Robust network design
VPN Conjecture
Oblivious routing

ABSTRACT

Robust network design refers to a class of optimization problems that occur when designing networks to efficiently handle variable demands. In this context, Fréchette et al. (2013) recently explored *hierarchical hubbing*: a routing strategy involving a multiplicity of "hubs" connected to terminals and each other in a treelike fashion. For a natural generalization of the VPN problem, we prove a structural characterization implying that the optimal hierarchical hubbing solution can be found efficiently, and relate this to a "Generalized VPN Conjecture".

1. Introduction

1.1. Robust network design

Robust network design considers the problem of building networks under uncertainty in the pattern of utilization. Introduced by Ben-Ameur and Kerivin [1], the framework encompasses the important case of the "hose model" introduced by Fingerhut [5] and Duffield et al. [3]. It can itself be seen as falling under the broader umbrella of robust optimization [2].

We refer the reader to [11] for a more in-depth treatment; here we will give a brief self-contained exposition of the model. We are given an undirected graph $G = (V, E)$; this should be interpreted as an existing high-capacity network, in which we can reserve capacity. We assume there is an unlimited total capacity on any given link of the network, and that the cost to reserve capacity on any link is a linear function of the capacity required. Let $c : E \to \mathbb{R}_+$ denote the per-unit cost of capacity on each edge. A set $W \subseteq V$ of *terminals* needs to be adequately connected using the capacity reserved.

A *traffic pattern* (or *demand pattern*) describes the precise pairwise demand requirements at some moment in time. It can be specified by a traffic matrix D, indexed by pairs of terminals; for terminals i, j, the entry D_{ij} represents the bandwidth needed to send data from i to j. In our network, the traffic pattern is not fixed, but varying (and possibly uncertain). To deal with this, the robust network design framework allows for a *set* of traffic patterns to be prescribed. This (it turns out) can always be taken to be a convex set, and so we describe this set, or *demand universe*, as a convex body $\mathcal{U} \subset \mathbb{R}_+^{W\times W}$.

In this paper, we will be concerned only with the case of *symmetric* demands, meaning that demand from i to j is not distinguishable from demand from j to i. In this case, it is convenient to consider $\mathcal{U}$ to be a subset of $\mathbb{R}_+^{\binom{W}{2}}$, where $\binom{W}{2}$ denotes the set of unordered pairs of terminals, so that $D_{ij} = D_{ji}$ refers to the same demand.

The robust network design (RND) problem asks for the cheapest capacity reservation $u : E \to \mathbb{R}_+$ that can support all traffic patterns in the specified universe $\mathcal{U}$. To fully specify the problem however, a further aspect must be considered: the *routing scheme*. The coarsest division is into *oblivious* or *dynamic* routing. In dynamic routing, the way in which traffic is routed may vary arbitrarily as a function of the current traffic pattern. This is typically infeasible, and we will be concerned here with the more practical oblivious routing, where the routing used for any given pair of terminals is specified in advance. We will also only consider *single-path routing*. The routing scheme in this case is described by a template $\mathcal{P} = \{P_{ij} : i, j \in W\}$, where P_{ij} is an i–j-path for each $i, j \in W$. (We do not require this path to be simple.) Since we consider symmetric demands, $P_{ij} = P_{ji}$ refers to the same path.

We may summarize the general robust network design problem (with oblivious, single-path routing and symmetric demands) as follows:

* Correspondence to: Department of Econometrics & Operations Research, Vrije Universiteit Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, Netherlands.
E-mail address: n.olver@vu.nl.

http://dx.doi.org/10.1016/j.orl.2015.12.020
0167-6377/

RND problem. *Given an undirected graph $G = (V, E)$ with edge costs $c(e)$, a terminal set $W \subseteq V$, and a convex demand universe $\mathcal{U} \subset \mathbb{R}_+^{\binom{W}{2}}$, a solution to the robust network design problem consists of a routing template $\mathcal{P} = \{P_{ij} : i, j \in W\}$, and a capacity allocation $u : E \to \mathbb{R}_+$, such that $\mathcal{U}$ can be routed according to $\mathcal{P}$ within the capacity u, i.e.,*

$$u(e) \geq \max_{D \in \mathcal{U}} \sum_{\{i,j\} \subseteq W} D_{ij} \ell(P_{ij}, e). \tag{1}$$

Here, $\ell(P, e)$ gives the number of times that edge e occurs on the (possibly non-simple) path P.

The difficulty in this optimization problem lies in choosing the routing template; once this is fixed, the optimal capacity allocation can be determined by solving a convex program described by (1), assuming we have access to at least a separation oracle for $\mathcal{U}$.

Note that there is always an optimum solution template whose paths P_{ij} are all simple, since any non-simple path can simply be replaced by a simple path within its support. The reason we allow non-simple paths is related to the specific type of routing templates we will be interested in.

The well-studied *symmetric hose model* [5,3] is parameterized by a vector $b \in \mathbb{R}_+^W$, yielding the universe

$$\mathcal{H}(b) = \left\{ D \in \mathbb{R}_+^{\binom{W}{2}} : \sum_{\{i,j\} \subset W} D_{ij} \leq b_i \quad \forall i \in W \right\}.$$

This models the situation where terminals are connected to the network with "hoses" of known, fixed capacity, so that the total demand involving terminal i cannot exceed the capacity b_i of its associated hose link. Any demand pattern that fits through the hoses should be routable in the final network. These hoses may model real links, or chosen based on operational criteria; either way, the hose model gives a simple, useful and concise description of what the network must be able to handle, making it a very popular model in the literature.

A number of variations and generalizations of this model have been considered in the literature [4,7,12,5]. For example, Fréchette et al. [7] consider the "capped" hose model, where in addition to the hose capacities b, point-to-point upper bounds on the demands are also given. In this paper, we will be interested in the following generalization introduced by Olver and Shepherd [12]. Let T^b be an arbitrary capacitated tree, with nonnegative edge capacities b and with leaf set in exact correspondence with the terminal set W. We will call any such capacitated tree a *demand tree*. We will use T^b to define a demand universe in a simple and natural way: let $\mathcal{U}(T^b)$ consist of all demand patterns that can be routed on T^b.

The case where T^b is a star corresponds precisely to the hose model; the capacity of the edge adjacent to terminal i precisely gives the marginal of i. This generalization allows the network operator more precise control over the demand universe, hopefully leading to more efficient solutions. In particular, if the terminals of the network can be logically divided into distinct groups (e.g., different branches of the company), with limited communication between groups, this information can be encoded via $\mathcal{U}(T^b)$.

We call the RND problem with oblivious routing for this class of demand universes the *generalized VPN problem*. It was shown in [12] that the generalized VPN problem is approximable to within a factor of 8.

1.2. Hierarchical hubbing

Fréchette et al. [7] define the following variant of the RND problem. Let T be a tree with leaf set W; we will call such a tree a *hub tree*. A T*-embedding* is a mapping of the internal nodes of T into the network, and a mapping of each edge e of T to a "cable" that connects the images of the endpoints of e (see Fig. 1). More than one node of T can be mapped to the same location, and multiple cables may run over the same edge of the network. More formally, we call a map $\varphi : V(T) \cup E(T) \to V(G) \cup E(G)$ a T-embedding if:

(i) $\varphi(v) \in V(G)$ for all $v \in V(T)$,
(ii) $\varphi(i) = i$ for all $i \in W$, and
(iii) $\varphi(vw)$ is a simple $\varphi(v)$–$\varphi(w)$ path in G for each $vw \in E(T)$.

The restriction to simple paths in the above definition is not necessary, but will be notationally convenient; in any case, there is no advantage to using non-simple paths. Such an embedding naturally defines a routing template: for each $\{i, j\} \subseteq W$, take the image of the unique i–j-path in the tree under the mapping (again, see Fig. 1), yielding an i–j-path in G. Note that this path need *not* be simple.

Given a hub tree T, along with a T-embedding, a T*-hubbing* solution is defined as follows. A T-embedding φ is chosen; this defines the routing template. Moreover, each cable is given a capacity, determined by the maximum load that can be placed on the cable by a demand in $\mathcal{U}$. Stated differently, for any given $f \in E(T)$, the cable associated with f is given capacity $b(f) := \max_{D \in \mathcal{U}} \sum_{i \in S, j \notin S} D_{ij}$, where S is the terminal set of either component of $T \setminus f$. The capacity $u(e)$ allocated to an edge $e \in E(G)$ must be at least the sum of the capacities of the cables running over that edge:

$$u(e) \geq \sum_{f \in E(T): e \in \varphi(f)} b(f) \quad \text{for all } e \in E(G).$$

The embedding φ and the valid capacity allocation u together describe the T-hubbing solution. A *hierarchical hubbing* solution is simply a T-hubbing solution, for some choice of a hub tree T. The cost of a hierarchical hubbing solution is defined simply as the cost of its associated capacity allocation.

So we have the following hierarchical hubbing RND problem (again, in the case of symmetric demands).

RND $_{\text{HH}}$ problem. *Given an undirected graph $G = (V, E)$ with edge costs $c(e)$, a terminal set $W \subseteq V$, and a convex demand universe $\mathcal{U} \subset \mathbb{R}_+^{\binom{W}{2}}$, the RND_{HH} problem is to find the tree T and an embedding of T that yields the cheapest hierarchical hubbing solution.*

Remark 1. It would also be natural to instead choose capacities by considering the routing template induced by the hierarchical hubbing, and using (1). This alternative formulation is in general not the same as described above; there may be situations where not all cables on a given edge can be simultaneously saturated by a traffic pattern in $\mathcal{U}$, leading to a larger capacity requirement with the cable formulation. The formulation that we use in this paper, and which is also used in [7], seems overall easier to deal with (e.g., see Lemma 2). If the Generalized VPN Conjecture discussed in Section 3 is true, it follows immediately that for the universe $\mathcal{U}(T^b)$, both formulations have a common optimal solution.

It is easy to confirm that any solution to the RND_{HH} problem is a solution to the RND problem, but not vice versa. So in general the optimal solution to RND_{HH} can be more expensive than the optimal RND solution; in fact, Fréchette et al. [7] demonstrate that the gap can be $\Omega(\log |V|)$, for some choices of the universe.

Fréchette et al. [7] are motivated to consider hierarchical hubbing for a few reasons. In *hub routing*, all traffic is routed via a single hub node; this has the advantage that routing decisions are localized at the hub. In order to address some practical shortfalls of hub routing, Shepherd and Winzer [13] ask for a "multihub" extension of this. Fréchette et al. argue that hierarchical hubbing provides a natural such extension (note that it is clearly a generalization; hub routing corresponds to taking the hub tree