An Improved Approximation Algorithm for Virtual Private Network Design

Friedrich Eisenbrand* Fabrizio Grandoni†

Abstract

Virtual private network design deals with the reservation of capacities in a network, such that the nodes can share communication. Each node in the network has associated upper bounds on the amount of flow that it can send to the network and receive from the network respectively. The problem then is to reserve capacities at minimum cost and to compute paths between every pair of nodes such that all valid traffic-matrices can be routed along the corresponding paths.

In this paper we present a simple 4.74-approximation algorithm for virtual private network design. The previous best approximation algorithm for this problem achieves a ratio of 5.55 (Gupta, Kumar, and Roughgarden STOC'03).

1 Introduction

Many problems in network design are built on the assumption that the *pairwise traffic* between the demand nodes is known in advance. Predicting this traffic however, is often illusive. The so-called *hose model* [6] allows greater flexibility. Here, each demand node has to know in advance an upper bound on the amount of flow that it wants to send into the network and an upper bound on the amount of flow that it wants to receive. *Virtual Private Network Design* (VPND) is the problem to reserve capacities at minimum cost, such that all possible pairwise flows, respecting the upper and lower bounds on the nodes, can be routed.

More formally, one is given an n-nodes undirected graph $G = (V,E)$ with nonnegative edge costs $c(e), e \in E$, a set of *demands* $D \subseteq V$, and for each demand $d \in D$ two upper bounds $b_{in}(d)$ and $b_{out}(d)$, which represent the maximum amount of flow that d can send and receive respectively. A *traffic matrix* T is a nonnegative $|D| \times |D|$ matrix, where $T(i,j)$ represents the flow from i to j. The traffic matrix T is called *valid* if it respects all upper and lower bounds on the demands, i.e., if

$$\sum_{j \in D, j \neq i} T(i,j) \leq b_{out}(i) \quad \text{and} \quad \sum_{j \in D, j \neq i} T(j,i) \leq b_{in}(i).$$

The problem is to compute capacities $u(e)$, $e \in E$, and paths $P_{(i,j)}$ for each ordered pair $(i,j) \in D \times D$ such that the next two conditions hold.

- All valid traffic-matrices can be routed along the corresponding paths.
- The total cost of the reservation $\sum_{e \in E} u(e)\, c(e)$ is minimal.

In this paper, following [9], we make some simplifying assumptions. By duplicating nodes, we can assume that each demand is either a *sender*, with $b^{out}_v = 1$ and $b^{in}_v = 0$, or a *receiver*, with $b^{out}_v = 0$ and $b^{in}_v = 1$. Let S and R be the set of senders and the set of receivers, respectively. The algorithms presented can be easily adapted such as to run in polynomial time even when the thresholds are not polynomially bounded. Moreover, by symmetry reasons, we can assume $|R| \geq |S|$.

Related work The problem was independently defined by Fingerhut et al. [7], and by Duffield et al. [6] and since then, studied by various authors. It follows from the results of Gupta et al. [8] that VPND is NP-complete. The authors provided a 2-approximation algorithm in the case that $b_{in}(j) = b_{out}(j)$ for each demand j. In this case, if the VPN is required to be a tree, the problem can also be solved optimally [8]. Oriolo [11] considers domination between traffic matrices. Gupta, Kumar and Roughgarden [9] provided the first constant factor approximation algorithm for VPND. Their algorithm achieves a guarantee of $(4+\rho)$ times the optimum, where ρ is the approximation ratio for the Steiner tree problem. Since $\rho \leq 1.55$ [12], the algorithm of Gupta et al. is a 5.55-approximation algorithm.

A related problem is *buy at bulk* network design (see e.g. [1, 2]). Here, there is a fixed demand $d_{i,j}$ between each pair of nodes in the graph, specifying the amount of flow which has to be sent from i to j. The costs of the capacities however is a concave function on the amount purchased, which reflects "economies of scale". Gupta et al. [9] consider the single source buy-at-bulk network design problem and present a simple constant factor approximation algorithm.

Another important issue in this context is to cope with arc failures in the network [3, 4]. Italiano, et al. [10] consider the problem of restoring the network, when at most one arc

*Max-Planck-Institut für Informatik, Stuhlsatzenhausweg 85, 66123 Saarbrücken, Germany, eisen@mpi-sb.mpg.de

†Max-Planck-Institut für Informatik, Stuhlsatzenhausweg 85, 66123 Saarbrücken, Germany, grandoni@mpi-sb.mpg.de

in a tree-solution to VPND might fail and provide a constant factor approximation algorithm.

Results In this paper we present a simple $(3+\sqrt{3}+o(1)) < 4.733$-approximation algorithm for VPND, thereby improving on the algorithm of Gupta et al.[9]. Our result is achieved by the following steps.

First, we present a very simple scheme, which provides a $2+|R|/|S|$-approximation. This is good if the number of receivers is small compared to the number of senders.

Second, we present a slightly refined analysis of the algorithm of Gupta, Kumar and Roughgarden [9] and show that their algorithm achieves a better ratio than 5.55 if the number of receivers is large compared to the number of senders. More precisely, we show that their algorithm computes an expected $3+\rho+2|S|/|R|$-approximation, where ρ is the approximation factor of the Steiner tree problem. Note that the first algorithm has an approximation guarantee of $3+\sqrt{3}$, as long as $|R|/|S| \leq 1+\sqrt{3}$, while the algorithm of Gupta et al. has this guarantee for $|R|/|S| \geq 2/(\sqrt{3}-\rho)$.

Third, we combine our simple $2+|R|/|S|$-approximation algorithm with a slight modification of the algorithm of Gupta et al. Here, a Steiner tree has to be computed, whose terminals are drawn from the receivers, each with probability $1/|S|$. If $|R|/|S|$ lies in the critical interval $1+\sqrt{3} < |R|/|S| < 2/(\sqrt{3}-\rho)$, then the expected number of terminals for the Steiner tree routine in the algorithm of Gupta et al. is a constant. This means that we can, most of the time, compute the *exact* Steiner tree solution in polynomial time. Our modification computes the exact Steiner tree as long as the number of terminals is logarithmic.

2 A simple $2+|R|/|S|$ approximation

Consider the following simple scheme: Choose a vertex $v \in V$ and reserve a unit of capacity along the shortest paths from v to each node in $R \cup S$. The effects of installing capacities along the shortest paths is cumulative. In other words, if k shortest paths share the same edge, we add k units of capacity to that edge. Also with a consistent tie-breaking rule, the edges with nonzero capacity form a tree. Such a solution to VPND is called a *tree-solution*.

It is clear that this reservation of capacity supports every valid traffic matrix. By $\ell(v,w)$ we denote the minimum distance between nodes v and w, w.r.t. the edge lengths $c(e)$. We argue now that there exists a sender $s' \in S$ with

$$\sum_{s\in S}\ell(s,s') + \sum_{r\in R}\ell(s',r) \leq (2+|R|/|S|)\,OPT, \tag{2.1}$$

where OPT is the cost of the optimal solution. This implies that the following simple algorithm is a $2+|R|/|S|$-approximation of VPND.

ALGORITHM 2.1. (VERY SIMPLE VPN)

1. Compute a vertex $v \in V$ such that $\sum_{s\in S}\ell(s,v) + \sum_{r\in R}\ell(v,r)$ is minimal.
2. Add one unit of capacity along the shortest path between each receiver $r \in R$ and v.
3. Add one unit of capacity along the shortest path between each sender $s \in S$ and v.

Similarly to the paper of Gupta et. al. [9], we consider the set of perfect matchings $\mathcal{M}$, which match the set of senders to a subset of the receivers. Given a matching $M \in \mathcal{M}$, one has $OPT \geq \sum_{(s,r)\in M}\ell(s,r)$, as a reservation has to support a path from s to r, for each $(s,r) \in M$ simultaneously. The set $\mathcal{M}$ has cardinality $|R|!/(|R|-|S|)!$, from which one can conclude the inequality

$$\sum_{M\in\mathcal{M}}\sum_{(s,r)\in M}\ell(s,r) \leq |R|!/(|R|-|S|)!\,OPT.$$

Each pair (s,r) appears in exactly $(|R|-1)!/(|R|-|S|)!$ matchings, from which one can conclude

$$\sum_{(r,s)\in R\times S}\ell(s,r) = \sum_{s\in S}\sum_{r\in R}\ell(s,r) \leq |R|\,OPT.$$

This implies that the average value of the sum of the shortest paths from a sender $s \in S$ to the receivers is bounded by $|R|/|S|\,OPT$ and in particular that there exists a sender s' with

$$\sum_{r\in R}\ell(s',r) \leq |R|/|S|\,OPT. \tag{2.2}$$

We now show that $\sum_{s\in S}\ell(s',s) \leq 2\,OPT$ holds. Let $R' \subseteq R$ be a subset of the receivers such that $|R'| = |S|$ and $\sum_{r\in R'}\ell(s',r)$ is minimal. It follows from (2.2) that

$$\sum_{r\in R'}\ell(s',r) \leq OPT. \tag{2.3}$$

Now consider a perfect matching $M' \in \mathcal{M}$, which matches S with R'. Clearly $\sum_{(s,r)\in M'}\ell(s,r) \leq OPT$. Let $(s,r) \in M'$. The triangle inequality implies $\ell(s,s') \leq \ell(s,r)+\ell(r,s')$. Thus one has with (2.3)

$$\sum_{s\in S}\ell(s,s') \leq \sum_{(s,r)\in M'}\ell(s,r) + \sum_{r\in R'}\ell(r,s') \leq 2\,OPT. \tag{2.4}$$

The inequalities (2.2) and (2.4) imply (2.1). Thus we have shown the following theorem.

THEOREM 2.1. *Algorithm 1 is a $2+|R|/|S|$ approximation algorithm for* VPND.

3 The algorithm of Gupta, Kumar and Roughgarden

In this section we review the algorithm of Gupta, Kumar and Roughgarden [9] and present a slightly refined analysis. It turns out that their algorithm has an expected approximation guarantee of $(3+\rho+2|S|/|R|)$, which is better than the previously proved bound of $(4+\rho)$ if $|S|/|R| < 1/2$. Here ρ is the approximation ratio for the Steiner tree problem. Robins and Zelikovsky [12] have shown that $\rho \leq 1.55$.

In Section 2 we presented a $2+|R|/|S|$ approximation. The factors $2+|R|/|S|$ and $(3+\rho+2|S|/|R|)$ are equal for

$$|R|/|S| = \left(1+\rho+\sqrt{(1+\rho)^2+8}\right)/2 < 3.18.$$

Since $2+|R|/|S|$ is increasing in $|R|/|S|$ and $(3+\rho+2|S|/|R|)$ is decreasing in $|R|/|S|$, the minimum of both values will always be at most 5.18. Thus the analysis below shows that a combination (taking the minimum solution) of the algorithm in Section 2 with the algorithm of Gupta et al. has an expected approximation guarantee of 5.18, which is already an improvement compared to the 5.55 approximation ratio of the algorithm of Gupta et al. alone.

ALGORITHM 3.1. (GKR [9])

1. Select a sender s' uniformly at random. Mark each receiver with probability $1/|S|$. Let R' be the set of the marked receivers.
2. Compute a ρ-approximate Steiner tree T on $F = \{s'\} \cup R'$. Add $|S|$ units of capacity to each edge of T.
3. Add one unit of capacity along the shortest paths between each receiver $r \in R$ and F.
4. Add one unit of capacity along the shortest paths between each sender $s \in S$ and F.

In the following we use $\ell(v,W) = \min_{w\in W}\{\ell(v,w)\}$ to denote the minimum distance of a vertex $v \in V$ to a subset $W \subseteq V$ of the vertices.

LEMMA 3.1. ([9]) *The expected cost* $E\left[c(T^*)\right]$ *of the optimum Steiner tree* T^* *with terminal set* F *satisfies*

$$E\left[c(T^*)\right] \leq OPT/|S|$$

and the total cost of the shortest paths from R *to* F *satisfies*

$$E\Big[\sum_{r\in R}\ell(r,F)\Big] = E\Big[\sum_{r\in R\setminus R'}\ell(r,F)\Big] \leq 2\,OPT.$$

We now come to our refined analysis.

THEOREM 3.1. *Algorithm 3.1 is an expected* $(3+\rho+2|S|/|R|)$*-approximation algorithm for* VPND.

Proof. Let APX denote the expected cost of the solution computed. One has

$$APX = E\Big[|S|\,c(T) + \sum_{r\in R}\ell(r,F) + \sum_{s\in S}\ell(s,F)\Big],$$

where $c(T)$ is the cost of the Steiner tree T computed in step (2). By Lemma 3.1 one has

(3.5)
$$E\big[|S|c(T)\big] \leq \rho\, OPT \quad \text{and} \quad E\Big[\sum_{r\in R}\ell(r,F)\Big] \leq 2\,OPT.$$

It remains to show that $E\left[\sum_{s\in S}\ell(s,F)\right] \leq (1+2\,|S|/|R|)OPT$ holds. If $|S| \leq |R'|$, then one has $\sum_{s\in S}\ell(s,F) \leq OPT$, from a matching argument as in Section 2.

Otherwise consider a subset $R'' \subseteq R\setminus R'$ of the receivers of cardinality $|S|-|R'|$ such that $\sum_{r\in R''}\ell(r,F)$ is minimal. Then one has

$$\sum_{r\in R''}\ell(r,F) \leq \frac{|S|-|R'|}{|R|-|R'|}\sum_{r\in R\setminus R'}\ell(r,F) \leq \frac{|S|}{|R|}\sum_{r\in R\setminus R'}\ell(r,F),$$

and together with Lemma 3.1 one has

$$E\Big[\sum_{r\in R''}\ell(r,F)\Big] \leq 2\,\frac{|S|}{|R|}\,OPT. \tag{3.6}$$

Let M be an arbitrary perfect matching between S and $R' \cup R''$. Since the cost of the matching is a lower bound on OPT, one has by the triangle inequality

$$\begin{aligned}\sum_{s\in S}\ell(s,F) &\leq \sum_{(s,r)\in M}\ell(s,r) + \sum_{r\in R'\cup R''}\ell(r,F)\\ &\leq OPT + \sum_{r\in R''}\ell(r,F).\end{aligned}$$

It follows then with (3.6) that

$$E\Big[\sum_{s\in S}\ell(s,F)\Big] \leq (1+2\,|S|/|R|)\,OPT \tag{3.7}$$

holds. The claim follows from (3.5) and (3.7).

At this point, we can already conclude a new structural result on the approximation of an optimal tree-solution to an optimal (graph) solution. Gupta et al.[9] could prove that every instance of VPND admits a tree solution with cost no more than 5 times the cost of an optimum (graph) solution. We can improve this factor.

COROLLARY 3.1. *Every instance of* VPND *admits a tree solution of cost at most* $(3+\sqrt{3})$ *times that of an optimum (graph) solution.*

Proof. Both algorithms, Algorithm 2.1 and Algorithm 3.1 compute tree solutions. The solution of Algorithm 2.1 is a $2+|R|/|S|$ approximation. If one computes the optimum Steiner tree in step (2) of Algorithm 3.1 one can replace ρ by 1 which yields a $4+2|S|/|R|$ approximation. Both values are equal if $|R|/|S| = 1+\sqrt{3}$, which implies that the minimum of both approximation ratios is always at most $3+\sqrt{3}$.

4 A refined combination

The proof of Corollary 3.1 shows that one could achieve a $3+\sqrt{3}$ approximation in polynomial time, if one could compute the optimal Steiner tree in step (2) efficiently. In this section, we show that we can get arbitrarily close to this bound with a polynomial algorithm.

Algorithm 2.1 has an approximation guarantee of $3+\sqrt{3}$, as long as $|R|/|S| \leq 1+\sqrt{3}$. Algorithm 3.1 has this guarantee for $|R|/|S| \geq 2/(\sqrt{3}-\rho)$. The expected approximation factor of the combination of both algorithms could be worse than $(3+\sqrt{3})$, if $1+\sqrt{3} < |R|/|S| < 2/(\sqrt{3}-\rho)$. The crucial observation now is that in this case the expected number of terminals of the Steiner tree problem occurring in step (2) of Algorithm 3.1 is at most $2/(\sqrt{3}-\rho)$. We denote this constant now by k.

An optimal Steiner tree on a graph with n nodes and t terminals can be computed in $O(3^t n + 2^t n^2 + n^3)$ time [5] with the Dreyfus-Wagner algorithm. This suggests the following variant of Algorithm 3.1, which computes an optimal Steiner tree, whenever $|R'| \leq k \log n$, where n is the number of nodes in G.

ALGORITHM 4.1. (MODIFIED GKR)

1. Select a sender s' uniformly at random. Mark each receiver with probability $1/|S|$. Let R' be the set of the marked receivers.
2. If $|R'| \leq k \log n$, compute the optimum Steiner tree T^* on F. Otherwise, compute a ρ-approximate Steiner tree on F. Let T be the tree computed. Add $|S|$ units of capacity to each edge of T.
3. Add one unit of capacity along the shortest paths between each receiver $r \in R$ and F.
4. Add one unit of capacity along the shortest paths between each sender $s \in S$ and F.

Clearly, Algorithm 4.1 is a polynomial time algorithm whose expected approximation guarantee is not worse than the one of Algorithm 3.1. What can be said about the approximation guarantee if $|R|/|S| \leq k$?

In that case, the expected size of R' is at most k. The probability, that the size of R' exceeds $\log n$ times its expected value is by Markov's inequality at most $1/\log n$. In this unlikely event however, we can estimate the outcome of the combination of both algorithms by the solution computed by Algorithm 2.1, which is a constant approximation algorithm in the case $|R|/|S| \leq k$. This is the intuition behind the proof of the next theorem.

THEOREM 4.1. *The combination (taking the cheaper solution) of Algorithm 2.1 and Algorithm 4.1 is an expected* $(3+\sqrt{3}+o(1))$*-approximation algorithm for* VPND.

Proof. Following the discussion above, we can restrict our analysis to the case where $1+\sqrt{3} < |R|/|S| < 2/(\sqrt{3}-\rho) = k$. Let APX denote the expected cost of the solution computed. One has, following the notation of the proof of Theorem 3.1,

$$APX = E\big[\min\{(2+|R|/|S|)\,OPT,\; |S|c(T) + \sum_{r\in R}\ell(r,F) + \sum_{s\in S}\ell(s,F)\}\big]. \tag{4.8}$$

Moreover,

$$\begin{aligned} APX &\leq E\Big[\sum_{r\in R}\ell(r,F) + \sum_{s\in S}\ell(s,F) \\ &\quad + \min\{(2+k)\,OPT, |S|c(T)\}\Big] \\ &\leq (3+2|S|/|R|)\,OPT \\ &\quad + E\big[\min\{(2+k)\,OPT, |S|\,c(T)\}\big]. \end{aligned}$$

Let A denote the event that $|R'| \leq k \log n$. By elementary probability theory one has

$$\begin{aligned} &E\big[\min\{(2+k)\,OPT, |S|c(T)\}\big] \\ &\leq P(A)\,E\big[|S|c(T) \mid A\big] + P(\overline{A})\,E\big[(2+k)\,OPT \mid \overline{A}\big]. \end{aligned} \tag{4.9}$$

We now consider both terms separately. By Markov's inequality one has $P(\overline{A}) \leq 1/\log n$. Thus

$$P(\overline{A})E\big[(2+k)\,OPT \mid \overline{A}\big] \leq (1/\log n)(2+k)OPT. \tag{4.10}$$

Given A, Algorithm 4.1 computes an optimal Steiner tree T^*. Also $E\big[c(T^*) \mid |R'| \leq h\big]$ is a non-decreasing function of h. Thus the first term on the right of (4.9) can be bounded by

$$P(A)\,E\big[|S|\,c(T) \mid A\big] \leq E\big[|S|\,c(T^*)\big] \leq OPT.$$

One therefore has

$$E\big[\min\{(2+k)\,OPT, |S|c(T)\}\big] \leq OPT(1+(2+k)/\log n), \tag{4.11}$$

from which we conclude that

$$APX \leq min\{2+|R|/|S|, 4+2|S|/|R|+(2+k)/\log n\}\,OPT. \tag{4.12}$$

Thus for each $\varepsilon > 0$, with increasing n, $APX \leq (3+\sqrt{3}+\varepsilon)$.

The above theorem shows that, for each $\varepsilon > 0$ the combination of Algorithm 2.1 and Algorithm 4.1 has an expected approximation ratio of $(3+\sqrt{3}+\varepsilon)$, if the number of nodes of the graph is sufficiently large. Since $4.733 > 3+\sqrt{3}$ we obtain the following corollary.

COROLLARY 4.1. *There exists a polynomial algorithm for* VPND *whose expected approximation ratio is at most* 4.733.

Acknowledgement We would like to thank Amit Kumar for many useful discussions and comments on this paper.

References

[1] M. Andrews and L. Zhang. The access network design problem. In *Proceedings of the 39th Annual Symposium on Foundations of Computer Science(FOCS-98)*, pages 40–49, Los Alamitos, CA, Nov.8–11 1998. IEEE Computer Society.

[2] B. Awerbuch and Y. Azar. Buy-at-bulk network design. In *38th Annual Symposium on Foundations of Computer Science*, pages 542–547, Miami Beach, Florida, 20–22 Oct. 1997. IEEE.

[3] G. Brightwell, G. Oriolo, and F. B. Shepherd. Reserving resilient capacity in a network. *SIAM J. Discrete Math.*, 14(4):524–539, 2001.

[4] G. Brightwell, G. Oriolo, and F. B. Shepherd. Reserving resilient capacity for a single commodity with upper-bound constraints. *Networks*, 41(2):87–96, 2003.

[5] S. E. Dreyfus and R. A. Wagner. The Steiner problem in graphs. *Networks*, 1:195–207, 1971/72.

[6] N. G. Duffield, P. Goyal, A. Greenberg, P. Mishra, K. K. Ramakrishnan, and J. E. van der Merive. A flexible model for resource management in virtual private networks. In *Proceedings of the conference on Applications, technologies, architectures, and protocols for computer communication*, pages 95–108. ACM Press, 1999.

[7] J. A. Fingerhut, S. Suri, and J. S. Turner. Designing least-cost nonblocking broadband networks. *Journal of Algorithms*, 24(2):287–309, 1997.

[8] A. Gupta, J. Kleinberg, A. Kumar, R. Rastogi, and B. Yener. Provisioning a virtual private network: a network design problem for multicommodity flow. In *ACM Symposium on the Theory of Computing*, pages 389–398, 2001.

[9] A. Gupta, A. Kumar, and T. Roughgarden. Simpler and better approximation algorithms for network design. In *ACM Symposium on the Theory of Computing*, pages 365–372, 2003.

[10] G. F. Italiano, R. Rastogi, and B. Yener. Restoration algorithms for virtual private networks in the hose model. In *Proceedings of the 21st Annual Joint Conference of the IEEE Computer and Communications Society (INFOCOM-02)*, volume 1 of *Proceedings IEEE INFOCOM 2002*, pages 131–139. IEEE Computer Society, 2002.

[11] G. Oriolo. Domination between traffic matrices. Technical report, Optimization Online, an e-print site for the optimization community, 2004.

[12] G. Robins and A. Zelikovsky. Improved steiner tree approximation in graphs. In *ACM-SIAM Symposium on Discrete Algorithms*, pages 770–779, 2000.