THEORY OF COMPUTING, Volume 8 (2012), pp. 401–413
www.theoryofcomputing.org

SPECIAL ISSUE IN HONOR OF RAJEEV MOTWANI

An $O(k^3 \log n)$-Approximation Algorithm for Vertex-Connectivity Survivable Network Design*

Julia Chuzhoy† Sanjeev Khanna‡

Received: August 13, 2010; published: August 17, 2012.

Abstract: In the Survivable Network Design Problem (SNDP), we are given an undirected graph $G(V,E)$ with costs on edges, along with an integer connectivity requirement $r(u,v)$ for each pair u,v of vertices. The goal is to find a minimum-cost subset E^* of edges such that in the subgraph of G induced by E^*, each pair u,v of vertices is $r(u,v)$-connected. In the *edge-connectivity version*, a pair u,v is $r(u,v)$-connected if there are $r(u,v)$ edge-disjoint paths between u and v. Similarly, in the *vertex-connectivity version*, a pair u,v is $r(u,v)$-connected if there are $r(u,v)$ vertex-disjoint paths between u and v. The edge-connectivity version of SNDP is known to have a factor 2 approximation. However, no non-trivial approximation algorithm has been known so far for the vertex version of SNDP, except for special cases of the problem.

We present an extremely simple randomized algorithm that achieves an $O(k^3 \log |T|)$-approximation for this problem, where k denotes the maximum connectivity requirement, and

*A preliminary version of this paper appeared in the 50th Annual IEEE Symposium on Foundations of Computer Science (FOCS), 2009.

†This research was supported in part by NSF CAREER award CCF-0844872 and Sloan Research Fellowship.

‡This research was supported in part by a Guggenheim Fellowship, an IBM Faculty Award, and by NSF Award CCF-0635084.

ACM Classification: F.2.2, G.1.6

AMS Classification: 68Q25, 68W20, 68W25

Key words and phrases: approximation algorithms, survivable network design, vertex-connectivity

DOI: 10.4086/toc.2012.v008a018

T is the set of vertices that participate in one or more pairs with non-zero connectivity requirements. We also give a simple proof of the recently discovered $O(k^2 \log |T|)$-approximation algorithm for the single-source version of vertex-connectivity SNDP. Our results establish a natural connection between vertex-connectivity and a well-understood generalization of edge-connectivity, namely, *element-connectivity*, in that an instance of vertex-connectivity SNDP can be reduced to a small number of instances of the element-connectivity SNDP.

1 Introduction

In the Survivable Network Design Problem (SNDP), we are given an undirected graph $G(V,E)$ with costs on edges, and an integer connectivity requirement $r(u,v)$ for each pair u,v of vertices. The goal is to find a minimum cost subset E^* of edges such that each pair (u,v) of vertices is connected by $r(u,v)$ paths in the subgraph induced by E^*. In the edge-connectivity version (EC-SNDP), these paths are required to be edge-disjoint, while in the vertex-connectivity version (VC-SNDP), they are required to be vertex-disjoint. It is not hard to show that EC-SNDP can be cast as a special case of VC-SNDP (see the full version of [8]). We denote by n the number of vertices in the graph and by k the maximum pairwise connectivity requirement, $k = \max_{u,v \in V} \{r(u,v)\}$. We also define a subset $T \subseteq V$ of vertices called *terminals*: a vertex $u \in T$ iff $r(u,v) > 0$ for some vertex $v \in V$.

Even for $k = 1$, both VC-SNDP and EC-SNDP are known to be APX-hard [2]; in fact the two problems are equivalent for $k = 1$, and this special case is commonly referred to as the minimum Steiner Forest problem.

1.1 Vertex-Connectivity SNDP

While a celebrated result of Jain [15] gives a 2-approximation algorithm for EC-SNDP, no non-trivial approximation algorithms are known for VC-SNDP, except for restricted special cases. Agrawal et al. [1] showed a 2-approximation algorithm for the special case when the maximum connectivity requirement $k = 1$, that is, the minimum Steiner Forest problem. For $k = 2$, a factor 2-approximation algorithm was given by Fleischer [11]. The k-vertex connected spanning subgraph problem, a special case of VC-SNDP where for all $u,v \in V$ $r(u,v) = k$, has been studied extensively. Cheriyan et al. [6, 7] gave an $O(\log k)$-approximation algorithm for this case when $k \le \sqrt{n/6}$, and an $O(\sqrt{n/\varepsilon})$-approximation algorithm for $k \le (1-\varepsilon)n$. For large k, Kortsarz and Nutov [18] improved the preceding bound to an $O(\log k \cdot \min\{\sqrt{k}, \log(k) \cdot n/(n-k)\})$-approximation. Fakcharoenphol and Laekhanukit [10] improved it to an $O(\log n \log k)$-approximation, and further obtained an $O(\log^2 k)$-approximation for $k < n/2$. Very recently, Nutov [22] obtained the best currently known approximation algorithm for the problem that gives an approximation ratio of $O(\log k \cdot \log(n/(n-k)))$.

Kortsarz et al. [17] showed that VC-SNDP is hard to approximate to within a factor of $2^{\log^{1-\varepsilon} n}$ for any $\varepsilon > 0$, unless $\mathsf{NP} \subseteq \mathsf{DTIME}\big(n^{\mathrm{polylog}(n)}\big)$, when k is polynomially large in n, that is, $k = \Omega(n^\delta)$ for some constant $\delta > 0$. This result was subsequently strengthened by Chakraborty et al. [3] to a k^ε-hardness for all $k > k_0$, where k_0 and ε are fixed positive constants. For constant k, their result holds under the assumption that $\mathsf{P} \neq \mathsf{NP}$, and for super-constant k, under the assumption that $\mathsf{NP} \not\subseteq \mathsf{DTIME}\big(n^{O(\log k)}\big)$. However, the existence of good approximation algorithms for small values of k has remained an open problem, even for

$k \geq 3$. In particular, when each connectivity requirement $r(u,v) \in \{0,3\}$, the best known approximation factor is polynomially large ($\tilde{O}(n)$ to best of our knowledge) while only an APX-hardness is known. The main result of our paper is an $O(k^3 \log |T|)$-approximation algorithm for VC-SNDP.

We note that subsequent to this work, Nutov [23] has obtained an $O(k^4 \log^2 |T|)$-approximation algorithm for the more general version of VC-SNDP, where the costs are on vertices (instead of edges), and the goal is to find a minimum cost subset V^* of vertices such that all vertex-connectivity requirements are satisfied in the subgraph induced by V^*.

Single-Source VC-SNDP A special case of VC-SNDP that has received much attention recently is the single-source version. In this problem there is a special vertex s called the *source*, and all non-zero connectivity requirements involve s, that is, if $u \neq s$ and $v \neq s$, then $r(u,v) = 0$. Kortsarz et al. [17] showed that even this restricted special case of VC-SNDP is hard to approximate up to factor $\Omega(\log n)$ unless $\mathsf{P} = \mathsf{NP}$, and recently Lando and Nutov [20] improved this to $(\log n)^{2-\varepsilon}$-hardness of approximation for any constant $\varepsilon > 0$, under the assumption that NP does not have quasi-polynomial time Las-Vegas algorithms. We note that both results only hold when k is polynomially large in n, that is, $k = \Omega(n^{\delta})$ for some constant $\delta > 0$. On the algorithmic side, Chakraborty et al. [3] gave an $k^{O(k^2)} \log^4 n$-approximation algorithm for the problem. This result was later independently improved to an $O(k^{O(k)} \log n)$-approximation by Chekuri and Korula [4], and to an $O(k^2 \log n)$-approximation by Chuzhoy and Khanna [8] and Nutov [24]. Subsequently, Chekuri and Korula [5] simplified the analysis of the algorithm of [8]. We note that for the uniform case, where all non-zero connectivity requirements are k, Chuzhoy and Khanna [8] show a slightly better $O(k \log n)$-approximation algorithm and the results of [5] extend to this special case. In this paper we give a simple $O(k^2 \log |T|)$-approximation algorithm for single-source VC-SNDP.

We note that subsequent to our work, Nutov [23] gave an $O(k \log k)$-approximation algorithm for single-source VC-SNDP. Nutov also obtains an improved approximation guarantee of $O(k^2 \log |T|)$ for single-source VC-SNDP when the costs are on vertices; the previous best known approximation ratio for this problem was $O(k^8 \log^2 n)$ [8].

1.2 Element-Connectivity SNDP

A closely related problem to EC-SNDP and VC-SNDP is the *element-connectivity* SNDP (Elem-SNDP). The input to Elem-SNDP is the same as for EC-SNDP and VC-SNDP. As before, we define the set $T \subseteq V$ of terminals to be vertices that participate in one or more pairs with a positive connectivity requirement. Given an Elem-SNDP instance, an *element* is any edge or any *non-terminal* vertex in the graph. We say that a pair s,t of vertices is *k-element connected* iff for every subset X of at most $(k-1)$ elements, s and t remain connected by a path when X is removed from the graph. In other words, there are k element-disjoint paths connecting s to t; these paths are allowed to share terminals. Another way to understand element-connectivity is to imagine that non-terminal vertices are placed on every edge, and a pair s,t is k-element connected iff there are k paths from s to t that do not share any non-terminal vertices. The goal in Elem-SNDP is to select a minimum-cost subset E^* of edges, such that in the graph induced by E^*, each pair u,v of vertices is $r(u,v)$-element connected.

Observe that in any graph, if a pair s,t is k-vertex connected, then it is also k-element connected, and similarly, if a pair s,t is k-element connected, then it is also k-edge connected. But the converse relationships do not hold, that is, if a pair s,t is k-edge connected, then it need not be k-element connected, and similarly, if a pair s,t is k-element connected, then it need not be k-vertex connected. Thus the notion of element-connectivity resides in between edge-connectivity and vertex-connectivity.

Elem-SNDP was introduced in [16] as a problem of intermediate difficulty between edge-connectivity and vertex-connectivity. Similar to edge-connectivity, element-connectivity is *transitive*, that is, for any three terminals t_1,t_2, and t_3, if t_1 and t_2 are k-element connected, and t_2 and t_3 are k-element connected, then t_1 and t_3 are k-element connected as well. Note that this transitivity property does not hold for vertex connectivity, as illustrated in Figure 1 for $k=2$.

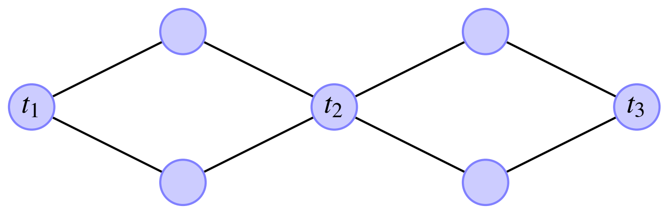

Figure 1: Each pair (t_i,t_j), $1 \le i < j \le 3$ above is 2-element connected (and hence 2-edge connected). The pairs (t_1,t_2) and (t_2,t_3) are 2-vertex connected but the pair (t_1,t_3) is only 1-vertex connected.

Jain et al. [16] give a primal-dual $O(\log k)$-approximation for this problem. Subsequently, Fleischer et al. [12] and Cheriyan et al. [7] gave a 2-approximation algorithm for Elem-SNDP via the iterative rounding technique, thus matching the 2-approximation guarantee of Jain [15] for EC-SNDP. We use the 2-approximation result for Elem-SNDP as a building block for our algorithms.

1.3 Our Results and Techniques

Our main result is as follows.

Theorem 1.1. *There is a polynomial-time randomized $O(k^3 \log|T|)$-approximation algorithm for VC-SNDP, where k is the largest pairwise connectivity requirement.*

The proof of this result is based on a randomized reduction that maps a given instance of VC-SNDP to a family of instances of Elem-SNDP. As noted earlier, the notion of element-connectivity is strictly weaker than vertex-connectivity. Our result shows that, nonetheless, an instance of VC-SNDP can be reduced to a collection of Elem-SNDP instances.

Specifically, given an instance of VC-SNDP, our reduction creates $O(k^3 \log|T|)$ instances of Elem-SNDP, where each instance is defined over an identical copy of the original graph, but with different connectivity requirements. The connectivity requirements for each instance are a carefully selected *subset* of the original connectivity requirements, and therefore, the value of the optimal solution to each resulting Elem-SNDP instance is bounded by the cost of the optimal solution to the original problem. With high probability, the reduction has the property that any collection of edges that is a feasible solution for *each one* of the Elem-SNDP instances generated above, is also a feasible solution for the given VC-SNDP instance. We can thus use the known 2-approximation algorithm for Elem-SNDP on each one of the $O(k^3 \log |T|)$ instances created, and combine them to obtain the desired result. We note here that our reduction also works for the more general variation of VC-SNDP, where the costs are on vertices instead of edges, and the goal is to select a minimum-cost subset V^* of vertices such that all connectivity requirements are satisfied in the subgraph induced by V^*.

Finally, we show that our ideas also yield a simple algorithm that gives an $O(k^2 \log |T|)$-approximation for the single-source VC-SNDP problem.

Organization: We present the proof of Theorem 1.1 in Section 2. Section 3 presents an alternative proof of the $O(k^2 \log |T|)$-approximation result for single-source VC-SNDP. In Section 4 we show a connection between our techniques and a well-studied notion in combinatorics and coding theory, namely, cover-free families. Using this connection we show that our bounds are essentially tight in that similar techniques cannot give significantly better approximation guarantees. Finally, we conclude in Section 5 with some directions for future work.

2 Algorithm for VC-SNDP

Recall that in VC-SNDP we are given an undirected graph $G(V,E)$ with costs on edges, and an integer connectivity requirement $r(u,v)$ for all $u,v \in V$. Additionally, we have a subset $T \subseteq V$ of terminals, and $r(u,v) > 0$ only if $u,v \in T$. The pairs of terminals with non-zero connectivity requirements are called *source-sink pairs*. The goal is to select a minimum-cost subset $E^* \subseteq E$ of edges, such that in the subgraph of G induced by E^*, denoted $G[E^*]$, every pair u,v is $r(u,v)$-vertex connected. We denote by k the maximum connectivity requirement, $k = \max_{u,v \in V} \{r(u,v)\}$, and we will use OPT to denote the cost of an optimal solution to the given VC-SNDP instance. Let (G,T,r) denote the input instance to VC-SNDP.

At a high-level, our algorithm is a polynomial-time randomized reduction from VC-SNDP to Elem-SNDP. Given an instance (G,T,r) of VC-SNDP, we create p identical copies of our input graph G, say $G_1, G_2, \ldots, G_p$, where p is a parameter to be specified later. For each copy G_i, we define a subset $T_i \subseteq T$ of terminals. We then view G_i as an instance of element-connectivity SNDP, where the connectivity requirements are induced by the set T_i of terminals as follows. For each $s,t \in T_i$ the new connectivity requirement is the same as the original one, and for all other pairs, the connectivity requirements are 0. Observe that for each G_i the cost of an optimal solution for the induced Elem-SNDP instance is at most OPT. We then apply the 2-approximation algorithm of [12, 7] to each one of the p instances of the k-element connectivity problem. Let E_i denote the set of edges output by the 2-approximation algorithm on the instance defined on G_i. Our final solution is $\tilde{E} = E_1 \cup E_2 \cup \cdots \cup E_p$. Since any solution to the original VC-SNDP instance is also a feasible solution for each one of the p element-connectivity instances created above, the cost of the solution above is bounded by $2p \cdot \mathrm{OPT}$.

We now show that for $p = O(k^3 \log |T|)$, there exist subsets $T_1, T_2, \ldots, T_p$ such that the solution $\tilde{E}$ produced above is a feasible solution for VC-SNDP. Moreover, we show a simple randomized algorithm to create the sets $T_1, T_2, \ldots, T_p$.

Definition 2.1. Given an instance (G,T,r) of VC-SNDP, let $\mathcal{M}$ be the set of the source-sink pairs. We say that a family $\{T_1, \ldots, T_p\}$ of subsets of T is *k-resilient* iff for each source-sink pair $(s,t) \in \mathcal{M}$, the following condition holds: for each subset $X \subseteq T \setminus \{s,t\}$ of size at most $(k-1)$, there is a subset T_i, $1 \le i \le p$, such that $s,t \in T_i$ and $X \cap T_i = \emptyset$.

We show below that a k-resilient family of subsets exists for $p = O(k^3 \log |T|)$, and give a poly-time randomized algorithm to find such a family with high probability. We start by proving that such a family guarantees that the algorithm produces a feasible solution.

Lemma 2.2. *Let $\{T_1,\dots,T_p\}$ be a k-resilient family of subsets. For each $1 \le i \le p$, let E_i be any feasible solution to the corresponding Elem-SNDP instance defined by T_i on G_i, and let $\tilde{E} = \bigcup_{i=1}^p E_i$. Then $\tilde{E}$ is a feasible solution to the input VC-SNDP instance.*

Proof. Let $(s,t) \in \mathcal{M}$ be any source-sink pair, and let $X \subseteq V \setminus \{s,t\}$ be any collection of at most $(r(s,t)-1) \le (k-1)$ vertices. It is enough to show that the removal of X from the graph $G[\tilde{E}]$ does not separate s from t. Let $X' = X \cap T$. Since $\{T_1,\dots,T_p\}$ is a k-resilient family of subsets, there is some T_i such that $s,t \in T_i$ while $T_i \cap X' = \emptyset$. Then X is a set of non-terminal vertices with respect to T_i. Recall that set E_i of edges defines a feasible solution to the element-connectivity SNDP instance corresponding to T_i. Since s is $r(s,t)$-element connected to t in the graph $G[E_i]$, the removal of set X from the graph $G[E_i]$ can not disconnect s from t. □

We now show how to construct a k-resilient family of subsets $\{T_1,\dots,T_p\}$. Let $p = 128k^3 \log |T|$, and set $q = p/(2k) = 64k^2 \log |T|$. Each terminal $t \in T$ selects q random indices uniformly and independently from the set $\{1,2,\dots,p\}$ (repetitions are allowed). Let $\phi(t)$ denote the set of indices chosen by the terminal t. For each $1 \le i \le p$, we then define $T_i = \{t \mid i \in \phi(t)\}$.

Lemma 2.3. *With high probability, the resulting family $\{T_1,\dots,T_p\}$ of subsets is k-resilient.*

Proof. We extend the definition of $\phi()$ to an arbitrary subset Z of vertices by defining

$$\phi(Z) = \bigcup_{t \in Z \cap T} \phi(t)\,.$$

Fix any source-sink pair (s,t). Let X be an arbitrary set of at most $(k-1)$ terminals that does not include s or t. Note that $|\phi(X)| \le (k-1)q < p/2$. We say that the *bad event* $\mathcal{E}_1(s,t,X)$ *occurs* if $|\phi(s) \cap \phi(X)| \ge 3q/4$. The expected value of $|\phi(s) \cap \phi(X)|$ is at most $q/2$ since at least half the indices in $\{1,2,\dots,p\}$ are disjoint from $\phi(X)$. Now using Chernoff bounds (see page 72 in [21])

$$\Pr[\mathcal{E}_1(s,t,X)] \le \Pr\left[|\phi(s) \cap \phi(X)| \ge \left(1+\frac{1}{2}\right)\frac{q}{2}\right] \le e^{-\frac{q}{2}\cdot(\frac{1}{2})^2\cdot\frac{1}{4}} \le e^{-\frac{q}{32}}\,.$$

We say that the *bad event* $\mathcal{E}_2(s,t,X)$ *occurs* if $\phi(s) \cap \phi(t) \subseteq \phi(X)$. We say that the set X is a *bad set* for a pair (s,t) if the event $\mathcal{E}_2(s,t,X)$ occurs. Note that if there is no bad set X of size at most $(k-1)$ for every pair $(s,t) \in \mathcal{M}$, then $\{T_1,\dots,T_p\}$ is a k-resilient family.

We observe that if event $\mathcal{E}_1(s,t,X)$ does not happen, then $|\phi(s) \setminus \phi(X)| \ge q/4$, so

$$\Pr\left[\mathcal{E}_2(s,t,X) \;\middle|\; \overline{\mathcal{E}_1(s,t,X)}\right] \le \left(1-\frac{q/4}{p}\right)^q \le e^{-\frac{q^2}{4p}} \le e^{-\frac{q}{8k}}\,.$$

Thus we can bound the probability of the event $\mathcal{E}_2(s,t,X)$ as follows:

$$\begin{aligned}
\Pr[\mathcal{E}_2(s,t,X)] &= \Pr[\mathcal{E}_2(s,t,X) \mid \mathcal{E}_1(s,t,X)] \Pr[\mathcal{E}_1(s,t,X)] + \Pr\left[\mathcal{E}_2(s,t,X) \;\middle|\; \overline{\mathcal{E}_1(s,t,X)}\right] \Pr\left[\overline{\mathcal{E}_1(s,t,X)}\right] \\
&\leq \Pr[\mathcal{E}_1(s,t,X)] + \Pr\left[\mathcal{E}_2(s,t,X) \;\middle|\; \overline{\mathcal{E}_1(s,t,X)}\right] \\
&\leq e^{-\frac{q}{32}} + e^{-\frac{q}{8k}} \\
&< |T|^{-4k}.
\end{aligned}$$

Hence, using the union bound, the probability that some bad set X of at most $(k-1)$ terminals exists for any pair (s,t) can be bounded by $|T|^{-2k}$. The lemma follows. □

Combining Lemmas 2.2 and 2.3, we thus obtain the following theorem.

Theorem 2.4. *Given a VC-SNDP instance (G,T,r), there is a polynomial-time randomized algorithm that creates $p = O(k^3 \log |T|)$ instances of Elem-SNDP, say $(G_1,T_1,r_1),(G_2,T_2,r_2),\ldots,(G_p,T_p,r_p)$, such that with high probability the following holds: if we are given, for each $1 \leq i \leq p$, any feasible solution E_i to the Elem-SNDP instance (G_i,T_i,r_i), then $\bigcup_{i=1}^{p} E_i$ is a feasible solution to the VC-SNDP instance (G,T,r). Moreover, for $1 \leq i \leq p$, the graph $G_i = G$, $T_i \subseteq T$, and the requirement function r_i is obtained by restricting r to T_i.*

Our main result follows as an easy corollary.

Corollary 2.5. *There is a randomized $O(k^3 \log |T|)$-approximation algorithm for VC-SNDP.*

Proof. With probability at least $1 - |T|^{-2k}$, the algorithm above generates a feasible solution for the given VC-SNDP instance. We can verify the feasibility of the generated solution by performing a max-flow computation for each source-sink pair. If the solution is not feasible, then we can use the trivial solution that solves optimally for each source-sink pair (a min-cost flow computation) and outputs union of all pairwise solutions. The cost of this trivial solution is at most $|T|^2 \cdot \mathrm{OPT}$. Thus the expected approximation ratio of the combined algorithm is bounded by

$$\left(1 - \frac{1}{|T|^{2k}}\right) O(k^3 \log |T|) + \left(\frac{1}{|T|^{2k}}\right) |T|^2 = O(k^3 \log |T|).$$

□

Remark 2.6. We note that this result implies that the standard set-pair relaxation for VC-SNDP [13] has an integrality gap of $O(k^3 \log |T|)$. This follows from the fact that the 2-approximation result of [12, 7] also establishes an upper bound of 2 on the integrality gap of the set-pair relaxation for element-connectivity. We also note that a lower bound of $\tilde{\Omega}(k^{1/3})$ is known on the integrality gap of the set-pair relaxation for VC-SNDP [3].

Remark 2.7. We also note that our reduction carries over to the node-weighted version of VC-SNDP, and in particular an α-approximation algorithm for the node-weighted element-connectivity SNDP implies an $O(\alpha k^3 \log |T|)$-approximation for the node-weighted VC-SNDP.

3 Algorithm for Single-Source VC-SNDP

In this section we show that an $O(k^2 \log |T|)$-approximation algorithm can be easily achieved using the above ideas for the single-source version of VC-SNDP.

Recall that the input to the single-source VC-SNDP is a graph $G(V,E)$ with a special vertex s called the source, and a subset T of terminals, where for each $t \in T$, we are given a connectivity requirement $r(s,t) \leq k$. We assume without loss of generality that $s \notin T$. The goal is to select a minimum-cost subset $E^* \subseteq E$ of edges, such that in the induced subgraph $G[E^*]$, every terminal $t \in T$ is $r(s,t)$-vertex connected to s. This is clearly a special case of VC-SNDP, where all source-sink pairs are of the form $\{(s,t)\}_{t \in T}$. As before, we create a family $\{T_1,\dots,T_p\}$ of subsets of terminals, $T_i \subseteq T$ for all $1 \leq i \leq p$. We also create p identical copies of our input graph G, say $G_1,\dots,G_p$. For each G_i we solve the single-source element-connectivity SNDP instance with connectivity requirements induced by terminals in T_i. Let E_i be the 2-approximate solution to instance G_i. Our final solution is $\tilde{E} = \bigcup_{i=1}^{p} E_i$. Clearly, the cost of the solution is at most $2p \cdot \mathrm{OPT}$.

Definition 3.1. A family $\{T_1,\dots,T_p\}$ of subsets of terminals is *weakly k-resilient* iff for each terminal $t \in T$, for each subset $X \subseteq T \setminus \{t\}$ of at most $(k-1)$ terminals, there is $i : 1 \leq i \leq p$, such that $t \in T_i$ and $X \cap T_i = \emptyset$.

Lemma 3.2. *Let $\{T_1,\dots,T_p\}$ be a weakly k-resilient family of subsets. For each $1 \leq i \leq p$, let E_i be any feasible solution to the corresponding Elem-SNDP instance defined by T_i on G_i, and let $\tilde{E} = \bigcup_{i=1}^{p} E_i$. Then $\tilde{E}$ is a feasible solution to the input VC-SNDP instance.*

Proof. Let $t \in T$ and let $X \subseteq V \setminus \{s,t\}$ be any subset of at most $r(s,t)-1 \leq (k-1)$ vertices excluding s and t. It is enough to prove that the removal of X from the graph $G[\tilde{E}]$ does not disconnect s from t. Let $X' = X \cap T$. Since $\{T_1,\dots,T_p\}$ is a weakly k-resilient family, there is some i, $1 \leq i \leq p$, such that $t \in T_i$ and $T_i \cap X' = \emptyset$. Consider the solution E_i to the corresponding k-element connectivity instance. Since vertices of X are non-terminal vertices for the instance G_i, their removal from the graph $G[E_i]$ does not disconnect s from t. □

We now show a randomized algorithm for constructing a weakly k-resilient family of subsets. Let $p = 4k^2 \log |T|$ and $q = p/(2k) = 2k \log |T|$. Each terminal $t \in T$ selects q indices from the set $\{1,2,\dots,p\}$ uniformly at random with repetitions. Let $\phi(t)$ denote the set of indices chosen by the terminal t. For each $1 \leq i \leq p$, we then define $T_i = \{t \mid i \in \phi(t)\}$.

Lemma 3.3. *With high probability, the resulting family of subsets $\{T_1,\dots,T_p\}$ is weakly k-resilient.*

Proof. Let $t \in T$ be any terminal and let X be any subset of at most $r(s,t)-1 \leq (k-1)$ terminals. As before, we extend the function ϕ to an arbitrary subset Z of vertices by defining $\phi(Z) = \bigcup_{t \in Z \cap T} \phi(t)$. We say that *bad event* $\mathcal{E}(t,X)$ *occurs* iff $\phi(t) \subseteq \phi(X)$. The probability of the event $\mathcal{E}(t,X)$ is at most

$$\left(1 - \frac{kq}{p}\right)^q = \left(\frac{1}{2}\right)^q \leq |T|^{-2k}.$$

Therefore, by union bound, the probability that the event $\mathcal{E}(t,X)$ happens for some terminal t and a set X of at most $(k-1)$ terminals, can be bounded by $|T|^{-k}$. The lemma follows. □

Combining Lemmas 3.2 and 3.3, we obtain the following corollary.

Corollary 3.4. *There is a randomized $O(k^2 \log |T|)$-approximation algorithm for single-source VC-SNDP.*

Proof. With probability at least $1 - |T|^{-k}$, the algorithm above generates a feasible solution for the given single-source VC-SNDP instance. We can verify the feasibility of the generated solution by performing a max-flow computation for each source-terminal pair. If the solution is not feasible, then we can use the trivial solution that solves optimally for each source-terminal pair (a min-cost flow computation) and outputs union of all pairwise solutions. The cost of this trivial solution is at most $|T| \cdot \mathrm{OPT}$. Thus the expected approximation ratio of the combined algorithm is bounded by

$$\left(1 - \frac{1}{|T|^k}\right) O(k^2 \log |T|) + \left(\frac{1}{|T|^k}\right) |T| = O(k^2 \log |T|)\,.$$

□

4 Resilient and cover-free families

The notion of a k-resilient and weakly k-resilient families is closely related to a well-studied notion in coding theory and combinatorics, namely, *cover-free* families of sets. A family $\mathcal{F}$ of sets over a universe $U = \{1, 2, \ldots, p\}$ is said to be *r-cover-free* if for all distinct $A, S_1, \ldots, S_r \in \mathcal{F}$, it satisfies the property that $A \not\subseteq \bigcup_{j=1}^{r} S_j$. This is precisely the property underlying our construction of a weakly k-resilient family. In particular, $\{T_1, T_2, \ldots, T_p\}$ is weakly k-resilient iff $\mathcal{F} = \{\phi(t) \mid t \in T\}$ is a $(k-1)$-cover-free family.

Let $N(r, \lambda)$ denote the smallest integer p such that there exists an r-cover-free family with λ sets over a universe of p elements. It is easy to see that the smaller the value $N(r, \lambda)$, the better the approximation guarantee achieved by the algorithm of Section 3. A classical result of Dyachkov and Rykov [9] (see the note by Füredi [14] for a simple proof of this lower bound result) shows that

$$N(r, \lambda) = \Omega\left(\frac{r^2 \log \lambda}{\log r}\right).$$

An immediate corollary of this result is that for any weakly k-resilient family for a set T of terminals, the parameter p must be

$$\Omega\left(\frac{k^2 \log |T|}{\log k}\right).$$

Thus the bound achieved by the simple randomized construction given in Lemma 3.3 is tight to within an $O(\log k)$ factor.

Kumar, Rajagopalan, and Sahai [19] gave an elegant deterministic construction for cover-free families based on Reed-Solomon codes. The construction gives slightly weaker guarantees than the randomized construction. For sake of completeness, we briefly describe their construction. Let $\mathbb{F}_q = \{u_1, u_2, \ldots, u_q\}$ be a finite field for some prime q. Moreover, let $F_{q,d}$ be the set of all polynomials over $\mathbb{F}_q$ of degree at most d where $d = q/k$. Consider the family of sets $\mathcal{F} = \{S_f \mid f \in F_{q,d+1}\}$ defined over the universe $U = \mathbb{F}_q \times \mathbb{F}_q$ where $S_f = \{\langle u_1, f(u_1)\rangle, \ldots, \langle u_q, f(u_q)\rangle\}$. Then $\mathcal{F}$ is a $(k-1)$-cover-free family since any two distinct polynomials in $F_{q,d}$ can agree on at most d points. Since the size of the underlying universe U is $p = q^2$ and $|\mathcal{F}| = \Omega(q^d)$, we get a deterministic construction for a weakly k-resilient family with

$$p = O\left(\frac{k^2 \log^2 |T|}{\log^2(k \log |T|)}\right).$$

A natural generalization of r-cover-free family is a (w,r)-cover-free family that is defined as follows. A family $\mathcal{F}$ of sets over a universe $U = \{1,2,\ldots,p\}$ is said to be (w,r)-*cover-free* if for all any $A_1, A_2, \ldots, A_w \in \mathcal{F}$ and any other $S_1, \ldots, S_r \in \mathcal{F}$, it satisfies the property that $\bigcap_{i=1}^{w} A_i \not\subseteq \bigcup_{j=1}^{r} S_j$. It is easy to see that $\{T_1, T_2, \ldots, T_p\}$ is k-resilient iff $\mathcal{F} = \{\phi(t) \mid t \in T\}$ is a $(2,k-1)$-cover-free family. Let $N(w,r,\lambda)$ denote the smallest integer p such that there exists a (w,r)-cover-free family with λ sets over a universe of p elements. Stinson, Wei, and Zhu [25] showed that for any $r \geq 1$, there exists a λ_0 that depends only on r, such that for all $\lambda \geq \lambda_0$

$$N(2,r,\lambda) = \Omega\left(\frac{r^3 \log \lambda}{\log r}\right).$$

An immediate corollary of this result is that for any k-resilient family for a set T of terminals, the parameter p must be

$$\Omega\left(\frac{k^3 \log |T|}{\log k}\right).$$

Thus the bound achieved by the simple randomized construction given in Lemma 2.3 is tight to within an $O(\log k)$ factor.

5 Conclusions

We presented a simple $O(k^3 \log |T|)$-approximation algorithm for VC-SNDP. Combined with the previously known $k^{\Omega(1)}$-hardness for this problem, the result presented here clarifies the approximability threshold of the general VC-SNDP problem. In particular, a polynomial factor dependence on k is both necessary and sufficient (modulo a logarithmic factor). An interesting question is whether VC-SNDP admits a constant factor approximation when the parameter k is a constant.

In contrast to the general VC-SNDP problem, there is a striking gap between upper and lower bounds for single-source VC-SNDP. In particular, when k is polynomially large in n, the best known hardness factor is $\Omega(\log^{2-\varepsilon} n)$ (for any constant ε) while the best currently achievable approximation ratio is $O(k \log k)$. It is an interesting open question to narrow this gap.

Acknowledgements

We thank Chandra Chekuri for his helpful comments on an earlier version of this paper. We are also grateful to anonymous referees whose comments have helped improve the presentation of this work.

References

[1] AJIT AGRAWAL, PHILIP KLEIN, AND R. RAVI: When trees collide: An approximation algorithm for the generalized Steiner problem on networks. *SIAM J. Comput.*, 24(3):440–456, 1995. Preliminary version in STOC'91. [doi:10.1137/S0097539792236237] 402

[2] MARSHALL BERN AND PAUL PLASSMANN: The Steiner problem with edge lengths 1 and 2. *Inform. Process. Lett.*, 32(4):171–176, 1989. [doi:10.1016/0020-0190(89)90039-2] 402

[3] TANMOY CHAKRABORTY, JULIA CHUZHOY, AND SANJEEV KHANNA: Network design for vertex connectivity. In *Proc. 40th STOC*, pp. 167–176. ACM Press, 2008. [doi:10.1145/1374376.1374403] 402, 403, 407

[4] CHANDRA CHEKURI AND NITISH KORULA: Single-sink network design with vertex connectivity requirements. In *IARCS Ann. Conf. on Foundations of Software Tech. and Theor. Comput. Sci. (FSTTCS'08)*, pp. 131–142. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2008. [doi:10.4230/LIPIcs.FSTTCS.2008.1747] 403

[5] CHANDRA CHEKURI AND NITISH KORULA: A graph reduction step preserving element-connectivity and applications. In *Proc. 36th Internat. Colloq. on Automata, Languages and Programming (ICALP'09)*, pp. 254–265. Springer, 2009. [doi:10.1007/978-3-642-02927-1_22] 403

[6] JOSEPH CHERIYAN, SANTOSH VEMPALA, AND ADRIAN VETTA: An approximation algorithm for the minimum-cost k-vertex connected subgraph. *SIAM J. Comput.*, 32(4):1050–1055, 2003. Preliminary version in STOC'02. [doi:10.1137/S0097539701392287] 402

[7] JOSEPH CHERIYAN, SANTOSH VEMPALA, AND ADRIAN VETTA: Network design via iterative rounding of setpair relaxations. *Combinatorica*, 26(3):255–275, 2006. [doi:10.1007/s00493-006-0016-z] 402, 404, 405, 407

[8] JULIA CHUZHOY AND SANJEEV KHANNA: Algorithms for single-source vertex connectivity. In *Proc. 49th FOCS*, pp. 105–114. IEEE Comp. Soc. Press, 2008. [doi:10.1109/FOCS.2008.63] 402, 403

[9] ARKADIĬ G. D'YACHKOV AND VYACHESLAV V. RYKOV: Bounds on the length of disjunctive codes. *Probl. Peredachi Inf.*, 18:7–13, 1982. [Math-Net.ru]. 409

[10] JITTAT FAKCHAROENPHOL AND BUNDIT LAEKHANUKIT: An $O(\log^2 k)$-approximation algorithm for the k-vertex connected spanning subgraph problem. In *Proc. 40th STOC*, pp. 153–158. ACM Press, 2008. [doi:10.1145/1374376.1374401] 402

[11] LISA FLEISCHER: A 2-approximation for minimum cost $\{0, 1, 2\}$ vertex connectivity. In *Proc. 8th Ann. Conf. on Integer Programming and Combinatorial Optimization (IPCO '01)*, pp. 115–129. Springer, 2001. [doi:10.1007/3-540-45535-3_10] 402

[12] LISA FLEISCHER, KAMAL JAIN, AND DAVID P. WILLIAMSON: Iterative rounding 2-approximation algorithms for minimum-cost vertex connectivity problems. *J. Comput. System Sci.*, 72(5):838–867, 2006. Preliminary version in FOCS'01. [doi:10.1016/j.jcss.2005.05.006] 404, 405, 407

[13] ANDRÁS FRANK AND TIBOR JORDÁN: Minimal edge-coverings of pairs of sets. *J. Combin. Theory Ser. B*, 65(1):73–110, 1995. [doi:10.1006/jctb.1995.1044] 407

[14] ZOLTÁN FÜREDI: On r-cover-free families. *J. Combin. Theory Ser. A*, 73(1):172–173, 1996. [doi:10.1006/jcta.1996.0012] 409

[15] KAMAL JAIN: A factor 2 approximation algorithm for the generalized Steiner network problem. *Combinatorica*, 21(1):39–60, 2001. Preliminary version in FOCS'98. [doi:10.1007/s004930170004] 402, 404

[16] KAMAL JAIN, ION MĂNDOIU, VIJAY V. VAZIRANI, AND DAVID P. WILLIAMSON: A primal-dual schema based approximation algorithm for the element connectivity problem. *J. Algorithms*, 45(1):1–15, 2002. Preliminary version in SODA'99. [doi:10.1016/S0196-6774(02)00222-5] 404

[17] GUY KORTSARZ, ROBERT KRAUTHGAMER, AND JAMES R. LEE: Hardness of approximation for vertex-connectivity network design problems. *SIAM J. Comput.*, 33(3):704–720, 2004. Preliminary version in APPROX'02. [doi:10.1137/S0097539702416736] 402, 403

[18] GUY KORTSARZ AND ZEEV NUTOV: Approximating k-node connected subgraphs via critical graphs. *SIAM J. Comput.*, 35(1):247–257, 2005. Preliminary version in STOC'04. [doi:10.1137/S0097539703435753] 402

[19] RAVI KUMAR, SRIDHAR RAJAGOPALAN, AND AMIT SAHAI: Coding constructions for blacklisting problems without computational assumptions. In *19th Ann. Internat. Cryptology Conf. (CRYPTO'99)*, pp. 609–623. Springer, 1999. [doi:10.1007/3-540-48405-1_38] 409

[20] YUVAL LANDO AND ZEEV NUTOV: Inapproximability of survivable networks. *Theoret. Comput. Sci.*, 410(21-23):2122–2125, 2009. Preliminary version in APPROX'08. [doi:10.1016/j.tcs.2009.01.036] 403

[21] RAJEEV MOTWANI AND PRABHAKAR RAGHAVAN: *Randomized Algorithms*. Cambridge University Press, 1995. 406

[22] ZEEV NUTOV: An almost $O(\log k)$-approximation for k-connected subgraphs. In *Proc. 20th Ann. ACM-SIAM Symp. on Discrete Algorithms (SODA'09)*, pp. 912–921. ACM Press, 2009. [ACM:1496770.1496869] 402

[23] ZEEV NUTOV: Approximating minimum cost connectivity problems via uncrossable bifamilies and spider-cover decompositions. In *Proc. 50th FOCS*, pp. 417–426. IEEE Comp. Soc. Press, 2009. [doi:10.1109/FOCS.2009.9] 403

[24] ZEEV NUTOV: A note on rooted survivable networks. *Inform. Process. Lett.*, 109(19):1114–1119, 2009. [doi:10.1016/j.ipl.2009.07.011] 403

[25] DOUGLAS R. STINSON, RUIZHONG WEI, AND LIE ZHU: Some new bounds for cover-free families. *J. Combin. Theory Ser. A*, 90(1):224–234, 2000. [doi:10.1006/jcta.1999.3036] 410

AUTHORS

Julia Chuzhoy
Toyota Technological Institute at Chicago
Chicago, IL 60637
cjulia@ttic.edu
http://ttic.uchicago.edu/~cjulia

Sanjeev Khanna
Dept. of Computer & Information Science
University of Pennsylvania
Philadelphia, PA 19104
sanjeev@cis.upenn.edu
http://www.cis.upenn.edu/~sanjeev

ABOUT THE AUTHORS

JULIA CHUZHOY is an Assistant Professor at the Toyota Technological Institute at Chicago. She received her Ph. D. from Technion-Israel Institute of Technology in 2004, under the supervision of Seffi Naor. She spent three years as a postdoc at MIT, the University of Pennsylvania, and the Institute for Advanced Study before joining TTIC in 2007. Her research area is approximation algorithms and hardness of approximation.

SANJEEV KHANNA is a Professor of Computer and Information Science and a Rosenbluth Faculty Fellow at University of Pennsylvania. He received a Ph. D. in Computer Science from Stanford University (1996) under the supervision of Rajeev Motwani, and undergraduate degrees in Computer Science and Economics from Birla Institute of Technology and Science, India (1990). From 1996 to 1999, he was a member of the Mathematical Sciences Research center at Bell Labs. He joined the University of Pennsylvania in 1999. Sanjeev's research interests are in algorithms and complexity with a focus on approximation algorithms and hardness of approximation.