Secluded Connectivity Problems

Shiri Chechik[1], Matthew P. Johnson[2] ⋆, Merav Parter[3] ⋆⋆, and David Peleg[3]

[1] Microsoft Research Silicon Valley Center, USA. schechik@microsoft.com
[2] Department of Electrical Engineering, UCLA, Los Angeles, USA. mpjohnson@gmail.com
[3] The Weizmann Institute of Science, Rehovot, Israel. {merav.parter,david.peleg}@weizmann.ac.il ⋆⋆⋆

Abstract. Consider a setting where possibly sensitive information sent over a path in a network is visible to every neighbor of (some node on) the path, thus including the nodes *on* the path itself. The *exposure* of a path P can be measured as the number of nodes adjacent to it, denoted by $N[P]$. A path is said to be *secluded* if its exposure is small. A similar measure can be applied to other connected subgraphs, such as Steiner trees connecting a given set of terminals. Such subgraphs may be relevant due to considerations of privacy, security or revenue maximization. This paper considers problems related to minimum exposure connectivity structures such as paths and Steiner trees. It is shown that on unweighted undirected n-node graphs, the problem of finding the minimum exposure path connecting a given pair of vertices is strongly inapproximable, i.e., hard to approximate within a factor of $O(2^{\log^{1-\epsilon} n})$ for any $\epsilon > 0$ (under an appropriate complexity assumption), but is approximable with ratio $\sqrt{\Delta} + 3$, where Δ is the maximum degree in the graph. One of our main results concerns the class of bounded-degree graphs, which is shown to exhibit the following interesting dichotomy. On the one hand, the minimum exposure path problem is NP-hard on *node-weighted* or *directed* bounded-degree graphs (even when the maximum degree is 4). On the other hand, we present a polynomial time algorithm (based on a nontrivial dynamic program) for the problem on unweighted undirected bounded-degree graphs. Likewise, the problem is shown to be polynomial also for the class of (weighted or unweighted)

⋆ The research of the second author was sponsored by the Army Research Laboratory and was accomplished under Cooperative Agreement Number W911NF-09-2-0053. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation here on.

⋆⋆ Recipient of the Google Europe Fellowship in distributed computing; research supported in part by this Fellowship.

⋆⋆⋆ Supported in part by the Israel Science Foundation (grant 894/09), United States-Israel Binational Science Foundation (grant 2008348), the I-CORE program of the Israel PBC and ISF (grant 4/11), Israel Ministry of Science and Technology (infrastructures grant), and Citi Foundation.

bounded-treewidth graphs. Turning to the more general problem of finding a minimum exposure Steiner tree connecting a given set of k terminals, the picture becomes more involved. In undirected unweighted graphs with unbounded degree, we present an approximation algorithm with ratio $\min\{\Delta, n/k, \sqrt{2n}, O(\log k \cdot (k + \sqrt{\Delta}))\}$. On unweighted undirected bounded-degree graphs, the problem is still polynomial when the number of terminals is fixed, but if the number of terminals is arbitrary, then the problem becomes NP-hard again.

1 Introduction

The problem. Consider a setting where possibly sensitive information sent over a path in a network is visible to every neighbor of (some node on) the path, thus including the nodes *on* the path itself. The *exposure* of a path P can be measured as the size (possibly node-weighted) of its *neighborhood* in this sense, denoted by $N[P]$. A path is said to be *secluded* if its exposure is small. A similar measure can be applied to other connected subgraphs, such as Steiner trees connecting a given set of terminals. Our interest is in finding connectivity structures with exposure as low as possible. This may be motivated by the fact that in real-life applications, a connectivity structure operates normally as part of the entire network G (and is not "extracted" from it), and so controlling the effect of its operation on the other nodes in the network may be of interest, in situations in which any "activation" of a node (by taking it as part of the structure) leads to an activation of its neighbors as well. In such settings, to minimize the set of total active nodes, we aim toward finding secluded or sufficiently private connectivity structures. Such subgraphs may be important in contexts where privacy is an important concern, or in settings where security measures must be installed on any node from which the information is visible, making it desirable to minimize their number. Another context where minimizing exposure may be desirable is when the information transferred among the participants has commercial value and overexposure to "free viewers" implies revenue loss.

This paper considers the problem of minimizing the exposure of subgraphs that satisfy some desired connectivity requirements. Two fundamental connectivity problems are considered, namely, single-path connectivity and Steiner trees, formulated as the `Secluded Path` and `Secluded Steiner Tree` problems, respectively, as follows. Given a graph $G = (V, E)$ and an s, t pair (respectively, a terminal set $\mathcal{S}$), it is required to find an $s - t$ path (respectively, a Steiner tree) of minimum exposure.

Related Work. The problems considered in this paper are variations of the classical shortest path and Steiner tree problems. In the standard versions of these problems, a cost measure is associated with edges or vertices, e.g., representing length or weight and the task is to identify a minimum cost subgraph satisfying the relevant connectivity requirement. Essentially, the cost of the solution subgraph is a *linear* sum of the solution's *constituent parts*, i.e., the sum of the weights of the edges or vertices chosen.

In contrast, in the setting of *labeled connectivity* problems, edges (and occasionally vertices) are associated with *labels* (or *colors*) and the objective is to identify a subgraph $G' \subseteq G$ that satisfies the connectivity requirements while minimizing the number of used labels. In other words, costs are now assigned to labels rather than to single edges. Such labeling schemes incorporate grouping constraints, based on partitioning the set of available edges into classes, each of which can be purchased in its entirety or not at all. These grouping constraints are motivated by applications from telecommunication networks, electrical networks, and multi-modal transportation networks. Labeled connectivity problems have been studies extensively from complexity-theoretic and algorithmic points of view [8,25,13,10]. The optimization problems in this category include, among others, the `Minimum Labeled Path` problem [13,25], the `Minimum Labeled Spanning Tree` problem [17,13], the `Minimum Labeled Cut` problem [26], and the `Labelled Prefect Matching` problem [22].

In both the traditional setting and the labeled connectivity setting, only edges or nodes that are explicitly part of the selected output structure are "paid for" in solution cost. That is, the cost of a candidate structure is a pure function of its components, ignoring the possible effects of "passive" participants, such as nodes that are "very close" to the structure in the input graph G. In contrast, in the setting considered in this paper, the cost of a connectivity structure G' is a function not only of its components but also of their immediate surroundings, namely, the manner in which G' is embedded in G plays a role as well. (Alternatively, we can say that the cost is a not necessarily a linear function of its components.)

A variant of the `Secluded Path` problem was recently introduced as the *Thinnest Path Problem* [11], where the focus was on directed hypergraph instances modeling transmission in wireless networks. In that application setting, each possible *transmission power* of a node yields a particular transmission range and hence a hyperedge directed from that node to the neighbors reached. The special case of that problem in which each node has a single possible transmission range is equivalent to our `Secluded Path` problem. They give a $\sqrt{\frac{n}{2}}$ approximation result for the `DegCost` algorithm (see Section 2) applied to this (more general) hypergraph setting.

The `Secluded Path` and `Secluded Steiner Tree` problems are related to several existing combinatorial optimization problems. These include the `Red-Blue Set Cover` problem [4,23], the `Minimum Labeled Path` problem [13,25] and the `Steiner Tree` [15] and `Node Weighted Steiner Tree` problems [16]. A prototypical example is the `Red-Blue Set Cover` problem, in which we are given a set R of red elements, a set B of blue elements and a family $S \subseteq 2^{|R|\cup|B|}$ of subsets of blue and red elements, and the objective is to find a subfamily $C \subseteq S$ covering all blue elements that minimizes the number of red elements covered. This problem is known to be strongly inapproximable.

Finally, turning to geometric settings, similarly motivated problems have been studied in the networking and sensor networks communities, where sensors are often modeled as unit disks. For example, the `Maximal Breach Path` problem [21] is defined in the context of traversing a region of the plane that contains sensor nodes at predetermined points, and its objective is to maximize the minimum distance between the points on the path and the sensor nodes. A dual problem studied extensively is *barrier coverage*, i.e., the (deterministic or stochastic) placement of sensors (see [18], [6]). Similarly motivated problems have been studied in the context of path planning in AI [14,19,20]. Although the motivation is similar, such problems are technically quite different from the graph-based problems studied here; those problems are typically posed in the geometric plane, amid obstacles that cause occlusion, and visibility is defined in terms of line-of-sight.

Contributions. In this paper, we introduce the concept of *secluded connectivity* and study some of its complexity and algorithmic aspects. We first state that the `Secluded Path` (and hence also `Secluded Steiner Tree`) problem is strongly inapproximable on unweighted undirected graphs with unbounded degree (more specifically, is hard to approximate with ratio $O(2^{\log^{1-\epsilon} n})$, where n is the number of nodes in the graph G, assuming $\mathcal{NP} \not\subseteq \mathcal{DTIME}\left(n^{\text{poly}\log n}\right)$). Conversely, we devise a $\sqrt{\Delta}+3$ approximation algorithm for the `Secluded Path` problem and a $\min\{\Delta, n/k, \sqrt{2n}, O(\log k\cdot(k+\sqrt{\Delta}))\}$ approximation algorithm for the `Secluded Steiner Tree` problem, where Δ is the maximum degree in the graph and k is the number of terminals.

One of our key results concerns bounded-degree graphs and reveals an interesting dichotomy. On the one hand, we show that `Secluded Path` is NP-hard on the class of *node-weighted* or *directed* bounded-degree graphs, even if the maximum degree is 4. In contrast, we show that on the class of unweighted undirected bounded-degree graphs, the `Secluded Path` problem admits an *exact* polynomial-time algorithm, which is based on a complex dynamic programming and requires some nontrivial analysis. Likewise, the `Secluded Steiner Tree` problem with fixed size terminal set is in P as well.

Finally, we consider some specific graph classes. We show that the `Secluded Path` and `Secluded Steiner Tree` problems are time polynomial for bounded-treewidth graphs. We also show that the `Secluded Path` (resp., `Secluded Steiner Tree`) problem can be approximated with ratio $O(1)$ (resp., $\Theta(\log k)$) in polynomial time for hereditary graph classes of bounded density. As an example, the `Secluded Path` problem has a 6 approximation on planar graphs. (A more careful direct analysis of the planar case yields ratio 3.)

2 Preliminaries and Notation

Consider a node-weighted graph $G(V, E, W)$, for some weight function $W : V \to \mathbb{R}_{\geq 0}$, with n nodes and maximum degree Δ. For a node $u \in V$, let $N(u) = \{v \in V \mid (u, v) \in E\}$ be the set of u's neighbors and let $N[u] = N(u) \cup \{u\}$ be u's *closed neighborhood*, i.e., including u itself. A *path* is a sequence $P = [u_1, \ldots, u_\ell]$, oriented from left to right, also termed a $u_1 - u_\ell$ path. Let $P[i] = u_i$ for $i \in \{1, \ldots, \ell\}$. Let $First(P) = u_1$ and $Last(P) = u_\ell$. For a path P and nodes x, y on it, let $P[x, y]$ be the subpath of P from x to y. For a connected subgraph

$G' \subseteq G$ and for $u_i, u_j \in V(G')$, let $\text{dist}_{G'}(u_i, u_j)$ be the distance between u_i and u_j in G'. Let $N(G') = \bigcup_{u \in G'} N(u) \setminus G'$ be the nodes that are strictly neighbors of G' nodes and $N[G'] = \bigcup_{u \in G'} N[u]$ be the set of nodes in the 1-neighborhood of G'. Define the cost of G' as

$$\texttt{Cost}(G') = \sum_{u \in N[G']} W(u) \ . \tag{1}$$

Note that if G is unweighted, then the cost of a subgraph G' is simply the cardinality of the set of G' nodes and their neighbors, $\texttt{Cost}(G') = |N[G']|$.

We sometimes consider the neighbors of node $u \in V(G)$ in different subgraphs. To avoid confusion, we denote $N_{G'}(u)$ the neighbors of u restricted to graph G'.

For a subgraph $G' \subseteq G$, let $\texttt{DegCost}(G')$ denote the sum of the degrees of the nodes of G'. If G' is a path, then this key parameter is closely related to our problem. It is not hard to see that for any given path P, $\texttt{Cost}(P) \leq \texttt{DegCost}(P)$. The problem of finding an $s-t$ path P with minimum $\texttt{DegCost}(P)$ is polynomial, making it a convenient starting-point for various heuristics for the problem.

In this paper we consider two main connectivity problems. In the `Secluded Path` problem we are given an unweighted graph $G(V, E)$, a source node s and target node t, and the objective is to find an $s - t$ path P with minimum neighborhood size. A generalization of this problem is the `Secluded Steiner Tree` problem, in which instead of two terminals s and t we are given a set of k terminal nodes $\mathcal{S}$ and it is required to find a tree T in G covering $\mathcal{S}$, of minimum neighborhood size. If the given graph G is node-weighted, then the *weighted* `Secluded Path` and `Secluded Steiner Tree` problems require minimizing the neighborhood cost as given in Eq. (1). We now define these tasks formally. For an $s - t$ pair, let $\mathcal{P}_{s,t} = \{P \mid P \text{ is a s-t path}\}$ be the set of all $s - t$ paths and let $q^*_{s,t} = \min\{\texttt{Cost}(P) \mid P \in \mathcal{P}_{s,t}\}$ be the minimum cost among these paths. Then the objective of the `Secluded Path` problem is to find a path $P^* \in \mathcal{P}_{s,t}$ that attains this minimum, i.e., such that $\texttt{Cost}(P^*) = q^*_{s,t}$. For the `Secluded Steiner Tree` problem, let $\mathcal{T}(\mathcal{S}) = \{T \subseteq G \mid \mathcal{S} \subseteq V(T), T \text{ is a tree}\}$ be the set of all trees in G covering $\mathcal{S}$, and let $q^*(\mathcal{S})$ be the minimum cost among these trees, i.e.,

$$q^*(\mathcal{S}) = \min\{\texttt{Cost}(T) \mid T \in \mathcal{T}(\mathcal{S})\} \ . \tag{2}$$

Then the solution for the problem is a tree $T^* \in \mathcal{T}(\mathcal{S})$ such that $\texttt{Cost}(T^*) = q^*(\mathcal{S})$. For paths P_1 and P_2, $P_1 \circ P_2$ denote the path obtained by concatenating P_2 to P_1.

3 Unweighted Undirected Graphs with Unbounded Degree

Hardness of approximation.

Theorem 1. *On unweighted undirected graphs with unbounded degree, the* `Secluded Path` *problem (and hence also the* `Secluded Steiner Tree` *problem)* *is strongly inapproximable. Specifically, unless* $\mathcal{NP} \subseteq \mathcal{DTIME}(n^{poly\log(n)})$*, the* `Secluded Path` *problem cannot be approximated to within a factor* $O(2^{\log^{1-\epsilon} n})$ *for any* $\epsilon > 0$*.*

Due to space limitation, missing proofs are provided in the full version of this paper [5].

Corollary 1. *The* `Secluded Path` *problem (and hence also the* `Secluded Steiner Tree` *problem) is strongly inapproximable in directed acyclic graphs.*

Approximation for the Secluded Path Problem

Theorem 2. *The* `Secluded Path` *problem in unweighted undirected graphs can be approximated within a ratio of* $\sqrt{\Delta}+3$*.*

Proof: Given an instance of the `Secluded Path` problem, let P^* be an $s-t$ path that minimizes $\texttt{Cost}(P^*)$. Note that we may assume without loss of generality that for every node u in P^*, the only neighbors of u in G from among the nodes of $V(P^*)$ are the nodes adjacent to u in P^*. To see this, note that otherwise, if u had an edge to some neighbor $u' \in V(P^*)$ such that e is not on P^*, we could have shortened the path P^* (by replacing the subpath from u to u' with the edge e) and obtained a shorter path with at most the same cost as P^*. Recall that $\texttt{DegCost}(P)$ denotes the sum of the node degrees of the path P, and that $\texttt{Cost}(P) \leq \texttt{DegCost}(P)$ for any P. Recall also that the problem of finding an $s-t$ path P with minimum $\texttt{DegCost}(P)$ is polynomial. We claim that the algorithm that returns the path Q^* minimizing $\texttt{DegCost}(Q^*)$ yields a $(\sqrt{\Delta}+3)$ approximation ratio for the `Secluded Path` problem. In order to prove this, we show that there exists a path Q such that $\texttt{DegCost}(Q) \leq (\sqrt{\Delta}+3) \cdot \texttt{Cost}(P^*)$. This implies that $\texttt{DegCost}(Q^*) \leq \texttt{DegCost}(Q) \leq (\sqrt{\Delta}+3) \cdot \texttt{Cost}(P^*)$, as required. The path Q is constructed by the following iterative process. Initially, all nodes are unmarked and we set $Q = P^*$. While there exists a node with more than $\sqrt{\Delta}+3$ unmarked neighbors on the path Q, pick such a node x. Let y be the first (closest to s) neighbor of x in Q and let z be the last (closest to t) neighbor of x in Q. Replace the subpath $Q[y,z]$ in Q with the path $[y,x,z]$ and mark the node x. We now show that $\texttt{DegCost}(Q) \leq (\sqrt{\Delta}+3) \cdot \texttt{Cost}(P^*)$. Let $X = \{x_1, \ldots x_\ell\}$ be the set of marked nodes in Q, where x_i is the node marked in iteration i. Note that $\ell = |X|$ is the number of iterations in the entire process. For the sake of analysis, partition $N[P^*]$ into two sets: S_1, those that have no unmarked neighbor in Q, and S_2, those that do. We claim that $|S_1| \geq \ell \cdot \sqrt{\Delta}$. For each $x_i \in X$, let $P(x_i)$ be the path that was replaced by the process of constructing Q in iteration i, and let $Q(x_i)$ be the path Q in the beginning of that iteration. Since x_i has more than $\sqrt{\Delta}+3$ unmarked neighbors in $Q(x_i)$, we get that $|P(x_i) \cap P^*| \geq \sqrt{\Delta}+4$. Let $P^-(x_i)$ be the path obtained by removing the first two nodes and last two nodes from $P(x_i)$. Note that $|P^-(x_i) \cap P^*| \geq \sqrt{\Delta}$. In addition, the sets $P^-(x_i) \cap P^*$ for $i = \{1, \ldots, \ell\}$ are pairwise disjoint, and moreover, none of the nodes in Q has a neighbor in $P^-(x_i) \cap P^*$. We thus get that $(P^-(x_i) \cap P^*) \subseteq S_1$. Note that $\texttt{Cost}(P^*) = |S_1| + |S_2|$ and that $\ell \leq |S_1|/\sqrt{\Delta}$.

We get that $\texttt{DegCost}(Q) \le |S_2| \cdot (\sqrt{\Delta}+3) + \ell \cdot \Delta \le |S_2| \cdot (\sqrt{\Delta}+3) + |S_1| \cdot \sqrt{\Delta} \le (|S_2| + |S_1|) \cdot (\sqrt{\Delta}+3) = \texttt{Cost}(P^*) \cdot (\sqrt{\Delta}+3)$. ∎

In addition, for the `Secluded Steiner Tree` problem with k terminals, in the full version [5], we establish the following.

Theorem 3. *The* `Secluded Steiner Tree` *problem in unweighted undirected graphs can be approximated within a ratio of* $\min\{\Delta, n/k, \sqrt{2n}, O(\log k \cdot (k + \sqrt{\Delta}))\}$.

4 Bounded-Degree Graphs

Weighted/Directed graphs. We first show that the `Secluded Path` problem (and thus also the `Secluded Steiner Tree` problem) is NP-hard on both directed graphs and weighted graphs, even if the maximum node degree is 4. We have the following.

Theorem 4. *The* `Secluded Path` *problem is NP-complete even for graphs of maximum degree* 4 *if they are either (a) node-weighted, or (b) directed.*

The unweighted undirected case. In contrast, for unweighted undirected case of bounded-degree graphs, we show that the `Secluded Path` problem is solvable in polynomial time.

Theorem 5. *The* `Secluded Path` *problem is solvable in* $O(n^2 \cdot \Delta^{\Delta+1})$ *time on unweighted undirected graphs, hence in* polynomial *time for degree-bounded graphs.*

Note that in the previous section we showed that if the graph is either weighted or directed then the `Secluded Path` problem (i.e., the special case of `Secluded Steiner Tree` with two terminals) is NP-hard. In addition, it is noteworthy that the related problem of `Minimum Labeled Path` problem [25,13] is NP-hard even for unweighted planar graph with max degree 4 (this follows from a straightforward reduction from `Vertex Cover`).

We begin with notation and couple of key observations in this context. For two subpaths P_1, P_2, define their asymmetric difference as

$$\text{Diff}(P_1, P_2) = |N[P_1] \setminus N[P_2]| \ .$$

Observation 6 *(a)* $\mathit{Diff}(P_1, P_2) \le \mathit{Diff}(P_1, P')$ *for every* $P' \subseteq P_2$.
(b) $\texttt{Cost}(P_1 \circ P_2) = \texttt{Cost}(P_1) + \mathit{Diff}(P_2, P_1)$.

For a given path P, let $\text{dist}_P(s, u)$ be the distance in edges between s and u in the path P. Recall, that Δ is the maximum degree in graph G.

Observation 7 *Let* u, v *be two nodes in some optimal* $s-t$ *path* P^* *that share a common neighbor, i.e.,* $N(u) \cap N(v) \neq \emptyset$. *Then* $\text{dist}_{P^*}(u, v) \le \Delta + 1$.

Proof: Assume for contradiction that there exists an optimal $s - t$ path P^* such that $N(u) \cap N(v) \neq \emptyset$ where $u = P^*[i]$ and $v = P^*[j]$, $1 < i < j$ for $(i - j) \geq \Delta + 2$. Recall that due to the optimality of P^*, for every node u in P^*, the only neighbors of u in G from among the nodes of $V(P^*)$ are the nodes adjacent to u in P^* (otherwise the path can be shortcut). Let $w \in N(u) \cap N(v)$ be the mutual neighbor of u and v and consider the alternative $s - t$ path $\widehat{P}$ obtained from P^* by replacing the subpath $Q = P[i, \ldots, j]$ by the subpath $P' = [u, w, v]$. Let $Q^- = P[i+2, \ldots, j-2]$ be an length-ℓ' internal subpath of Q where $\ell' \geq \Delta - 1$. Then since the degree of w is at most Δ, it follows that

$$\texttt{Cost}(\widehat{P}) \leq N[V(P) \setminus V(Q^-)] + \Delta - 2 \ , \tag{3}$$

where $\Delta - 2$ is an upper bound on the number $w's$ neighbors other than u and v. In addition, note that by the optimality of P^* it contains no shortcut and hence $V(Q^-) \cap (N[V(P) \setminus V(Q^-)]) = \emptyset$ (i.e., for node $P[i]$ the only neighbors on the path P are $P[i-1]$ and $P[i+1]$). Thus,

$$\begin{aligned} \texttt{Cost}(P^*) &\geq N[V(P) \setminus V(Q^-)] + |V(Q^-)| \\ &\geq N[V(P) \setminus V(Q^-)] + \Delta - 1 > \texttt{Cost}(\widehat{P}) \ , \end{aligned}$$

where the last inequality follows from Eq. (3), contradicting the optimality of P^*. The observation follows. ∎

In other words, the observation says that two vertices on the optimal path at distance $\Delta + 2$ (which is constant for bounded-degree graphs) or more have no common neighbors. This key observation is at the heart of our dynamic program, as it enables the necessary subproblem independence property. The difficulty is that this observation applies only to optimal paths, and so a delicate analysis is needed to justify why the dynamic program works. Note that the main difficulty of computing the optimal secluded path P^* is that the cost function $\texttt{Cost}(P^*)$ is not a linear function of path's components as in the related `DegCost` measure. Instead, the residual cost of the ith vertex in the path *depends* on the neighborhood of the length-$(i-1)$ prefix of P^*. This dependency implies that the secluded path computation cannot be simply decomposed into independent subtasks. However, in contrast to suboptimal $s-t$ paths, the dependency (due to mutual neighbors) between the components of an optimal path is *limited* by the maximum degree Δ of the graph. The *limited dependency* exhibited by any $s-t'$ optimal path facilitates the correctness of the dynamic programming approach. Essentially, in the dynamic program, entries that correspond to subsolution σ of *any* optimal $s - t'$ path enjoy the limited dependency, and hence the values computed for these entries correspond to the *exact* cost of an optimal path that starts with s and ends with σ. In contrast, entries of subpaths σ that do not participate in any optimal $s-t'$ path, correspond to an *upper bound* on the cost of some path P that starts with s and ends with σ. This is due to the fact that the value of the entry is computed under the limited dependency assumption, and thus does not take into account the possible double counting of mutual neighbors between *distant* vertices in the path. Therefore, the possibly "falsified" entries cannot compete with the exact values, which are guaranteed to be computed for the entries that hold the subpaths of the optimal path. This informal intuition is formalized below.

For a path P of length $\ell \geq \Delta + 1$, let $\text{Suff}(P) = \langle v_{\ell-\Delta}, \ldots, v_\ell \rangle$ be the $(\Delta + 1)$-suffix of P.

Lemma 1. *Let P^* be an optimal $s - t'$ path of length $\geq \Delta + 1$. Let $P^* = P_1 \circ P_2$ be some partition of P^* into two subpaths such that $|P_1| \geq \Delta + 1$. Then $\mathit{Diff}(P_2, P_1) = \mathit{Diff}(P_2, \text{Suff}(P_1))$.*

Proof: For ease of notation, let $\ell = |P_1|$, $P' = P_1[1, \ldots, \ell - (\Delta + 1)]$, and $\sigma = \text{Suff}(P_1)$. By definition, $\Delta(P_2, P_1) = |\,(N[P_2] \setminus N[\sigma]) \setminus N[P']|$. In the same manner, $\Delta(P_2, \sigma) = |\,(N[P_2] \setminus N[\sigma])\,|$. Assume for contradiction that the lemma does not hold, namely, $\Delta(P_2, \sigma) \neq \Delta(P_2, P' \circ \sigma)$. Then by Obs. 6(a) we have that $\Delta(P_2, \sigma) > \Delta(P_2, P' \circ \sigma)$. This implies that $N[P_2] \cap N[P'] \neq \emptyset$. Let $u \in N[P_2] \cap N[P']$. There are two cases to consider: (a) $u \in P'$, and (b) u has a neighbor v_1 in P'. We handle case (a) by further dividing it into two subcases: (a1) $u \in P_2$, i.e., u occurs at least twice in P^*, once in P' and once in P_2, and (a2) u has a neighbor v_2 in P_2. Note that in both subcases, there exists a shortcut of P^*, obtained in subcase (a1) by cutting the subpath between the two duplicates of u and in subcase (a2) by shortcutting from u to v_2. This shortcut results in a strictly lower cost path, in contradiction to the optimality of P^*. We proceed with case (b). Let $v_2 \in P_2$ be such that $u \in N[v_2]$. If $v_2 = u$, then clearly the path can be shortcut by going from $v_1 \in P'$ directly to $u \in P_2$, resulting in a lower cost path, in contradiction to the optimality of P^*. If $v_2 \neq u$, then $dist_{P^*}(v_1, v_2) \geq \Delta + 2$ (since $|\sigma| \geq \Delta + 1$) and $N[v_1] \cap N(v_2) \neq \emptyset$, which in contradiction to Obs. 7. The Lemma follows. ∎

Corollary 2. *Let P^* be an optimal $s - t'$ path of length $\geq \Delta + 1$. Let $P^* = P_1 \circ P_2$ be some partition of P^* into two subpaths such that $|P_1| \geq \Delta + 1$. Then $\texttt{Cost}(P^*) = \texttt{Cost}(P_1) + \mathit{Diff}(P_2, \text{Suff}(P_1))$.*

For clarity of representation, we describe a polynomial algorithm for the `Secluded Path` problem, i.e., where $\mathcal{S} = \{s, t\}$ and in addition $\Delta \leq 3$, in which case $\text{Suff}(P) = \langle v_{\ell-3}, \ldots, v_\ell \rangle$. The general case of $\Delta = O(1)$ is immediate by the description for the special case of $\Delta = 3$. The case of `Secluded Steiner Tree` with a fixed number of terminals $|\mathcal{S}| = O(1)$ is described in [5].

The algorithm we present is based on dynamic programming. For each $4 \leq i \leq n$, and every length-4 subpath given by a quartet of nodes $\sigma \subseteq V(G)$, it computes a length-i path $\pi(\sigma, i)$ that starts with s and ends with σ (if such exists) and an upper bound $f(\sigma, i)$ on the cost of this path, $f(\sigma, i) \geq \texttt{Cost}(\pi(\sigma, i))$. These values are computed inductively, using the values previously computed for other σ's and $i - 1$. In contrast to the general framework of dynamic programming, the interpretation of the computed values $f(\sigma, i)$ and $\pi(\sigma, i)$, namely, the relation between the dynamic programming values $f(\sigma, i)$ and $\pi(\sigma, i)$ and some "optimal" counterparts is more involved. In general, for arbitrary σ and i, the path $\pi(\sigma, i)$ is not guaranteed to be optimal in any sense and neither is its corresponding value $f(\sigma, i)$ (as $f(\sigma, i) \geq \texttt{Cost}(\pi(\sigma, i))$). However, quite interestingly, there is a subset of quartets for which a useful characterization of $f(\sigma, i)$ and $\pi(\sigma, i)$ can be established. Specifically, for every $4 \leq i \leq n$, there is a subclass of quartets Ψ_i^*, for which the computed values are in fact "optimal", in the sense that for every $\sigma \in \Psi_i^*$, the length-i path $\pi(\sigma, i)$ is of minimal cost among all other length-i paths that start with s and end with σ. We call such a path a *semi-optimal* path, since it is optimal only restricted to the length and specific suffix requirements. It turns out that for the special class of quartets Ψ_i^*, every semi-optimal path of $\sigma \in \Psi_i^*$ is also a prefix of some optimal $s - t'$ path. This property allows one to apply Cor. 2, which constitutes the key ingredient in our technique. In particular, it allows us to establish that $f(\sigma, i) = \texttt{Cost}(\pi(\sigma, i))$. This is sufficient for our purposes since the set $\bigcup_{i=1}^{n} \Psi_i^*$ contains *any* quartet that occurs in *some* optimal secluded $s - t$ path P^*. Specifically, for every $s - t$ optimal path P^*, it holds that the quartet $\mathrm{Suff}(P^*) = P^*[i-3, i]$ satisfies $\mathrm{Suff}(P^*) \in \Psi_i^*$. The correctness of the dynamic programming is established by the fact that the values computed for quartets that occur in optimal paths in fact correspond to optimal values as required (despite the fact the these values are "useless" for other quartets).

The algorithm. If the shortest path between s and t is less than 3, then the optimal `Secluded Path` can be found by an exhaustive search, so we assume throughout that $dist_G(s, t) \geq 3$. Let Ψ be the set of all length-4 subpaths σ in G. That is, for every $\sigma \in \Psi$, we have $V(\sigma) \subseteq V(G)$ and $(\sigma[i], \sigma[i+1]) \in E(G)$ for every $i \in [1, 3]$. For $\sigma \in \Psi$, define the collection of *shifted successor* of G as $Next(\sigma) = \{\langle \sigma[2], \ldots, \sigma[4], u\rangle \mid u \in (N[\sigma[4]] \setminus \sigma)\}$. For every pair (σ, i), where $\sigma \in \Psi$ and $i \in \{1, \ldots, n\}$, the algorithm computes a value $f(\sigma, i)$ and length-i path $\pi(\sigma, i)$ ending with σ (i.e., $\pi(\sigma, i)[i-3, \ldots, i] = \sigma$). These values are computed inductively. For $i = 4$, let $\pi(\sigma, 4) = \sigma$ and

$$f(\sigma, 4) \;=\; \begin{cases} \texttt{Cost}(\sigma), & \text{if } \sigma[1] = s \\ \infty, & \text{otherwise} \end{cases}$$

Once the algorithm has computed $f(\sigma, j)$ for every $4 \leq j \leq i-1$ and every $\sigma \in \Psi$, in step $i \in \{5, n\}$ it computes

$$f(\sigma, i) = \min\{f(\sigma', i-1) + \mathrm{Diff}(\sigma, \sigma') \;\mid\; \sigma \in Next(\sigma')\}. \tag{4}$$

Note that

$$\mathrm{Diff}(\sigma, \sigma') = \mathrm{Diff}(\sigma[4], \sigma') = |N[\sigma[4]] \setminus N[\sigma']|. \tag{5}$$

Let $\sigma' \in \Psi$ such that $\sigma \in Next(\sigma')$ and σ' achieves the minimum value in Eq. (4). Then define $Pred(\sigma) = \sigma'$ and let

$$\pi(\sigma, i) = \pi(Pred(\sigma), i-1) \circ \sigma[4]. \tag{6}$$

Let $q_{s,t}^* = \min\{f(\sigma, i) \mid i \in \{4, \ldots, n\}, Last(\pi(\sigma, i)) = t\}$, and set $P^* = \pi(\sigma^*, i^*)$ for σ^*, i^* such that $f(\sigma^*, i^*) = q_{s,t}^*$. Note that there are at most $O(n^2)$ entries $f(\sigma, i)$, each computed in constant time, and so the overall running time is $O(n^2)$. In [5], we provide a detailed analysis and establish Thm. 5. For the `Secluded Steiner Tree` problem, we show the following.

Theorem 8. *On unweighted undirected degree-bounded graphs, we have the following: (a) for arbitrary k, the* `Secluded Steiner Tree` *problem is NP-hard; (b) for $k = O(1)$, the* `Secluded Steiner Tree` *problem is* polynomial*.*

5 Secluded Connectivity for Specific Graph Families

Bounded-treewidth graphs. For a graph $G(V, E)$, let $TW(G)$ denote the *treewidth* of G. For bounded treewidth graph, i.e. $TW(G) = O(1)$, we have the following.

Theorem 9. *The* `Secluded Steiner Tree` *problem (and hence also the* `Secluded Path` *problem) can be solved in* linear *time for graphs with fixed treewidth. In addition, given the tree decomposition of G, the* `Secluded Steiner Tree` *problem is solvable in $\widetilde{O}(n^3)$ if $TW(G) = O(\log n / \log\log n)$. This holds even for weighted and directed graphs.*

Bounded Density Graphs. Let $\texttt{DegCost}^*(\mathcal{S}) = \min\{\texttt{DegCost}(T) \mid T \in \mathcal{T}(\mathcal{S})\}$. In the full version [5], we show the following.

Proposition 1. *Let $\mathcal{G}$ be a hereditary class of graphs with a linear number of edges, i.e., a set of graphs such that for each $G \in \mathcal{G}$, $|E(G)| \leq \ell \cdot |V(G)|$ for some constant ℓ, and where $G \in \mathcal{G}$ implies that $G' \in \mathcal{G}$ for each subgraph G' of G. Then $\texttt{DegCost}^*(\mathcal{S}) \leq 2\ell \cdot q^*(\mathcal{S})$.*

An example of such a class of graphs is the family of planar graphs. For this family, the above proposition yields a 6-approximation for the `Secluded Path` problem. (A more careful direct analysis for planar graphs yields a 3-approximation.) Finally, note that whereas computing $\texttt{DegCost}^*(\mathcal{S})$ for a constant number of terminals k is polynomial, for arbitrary k, computing $\texttt{DegCost}^*(\mathcal{S})$ is NP-hard but can be approximated to within a ratio of $\Theta(\log k)$, see [5]. Thus, for the class of bounded density graphs, by Prop. 1, the `Secluded Path` problem has a constant ratio approximation and the `Secluded Steiner Tree` problem has an $\Theta(\log k)$ ratio approximation.

Theorem 10. *For the class of bounded-density graphs, the* `Secluded Path` *problem (respectively,* `Secluded Steiner Tree` *problem) is approximated within a ratio of $O(1)$ (respectively, $\Theta(\log k)$).*

Acknowledgment We are grateful to Amotz Bar-Noy, Prithwish Basu and Michael Dinitz for helpful discussions.

References

1. H.L. Bodlaender. A linear time algorithm for finding tree-decompositions of small treewidth. In *STOC*, 226–234, 1993.
2. H.L. Bodlaender. A tourist guide through treewidth. *Acta Cybern.*, 11:1–22, 1993.

3. H.L. Bodlaender. Treewidth: Structure and algorithms. In *SIROCCO*, 11–25, 2007.
4. R.D. Carr, S. Doddi, G. Konjevod, and M.V. Marathe. On the red-blue set cover problem. In *SODA*, 345–353, 2000.
5. S. Chechik, M. P. Johnson, M. Parter, and D. Peleg Secluded Connectivity Problems. http://arxiv.org/abs/1212.6176.
6. A. Chen, S. Kumar, and T.-H. Lai. Local barrier coverage in wireless sensor networks. *IEEE Tr. Mob. Comput.*, 9:491–504, 2010.
7. M. Chimani, P. Mutzel, and B. Zey. Improved steiner tree algorithms for bounded treewidth. In *IWOCA*, 374–386, 2011.
8. I. Dinur and S. Safra. On the hardness of approximating label-cover. *IPL*, 89:247–254, 2004.
9. U. Feige. A threshold of ln n for approximating set cover. *J. ACM*, 45:634–652, 1998.
10. M.R. Fellows, J. Guo, and I.A. Kanj. The parameterized complexity of some minimum label problems. *JCSS*, 76:727–740, 2010.
11. J. Gao, Q. Zhao, and A. Swami. The Thinnest Path Problem for Secure Communications: A Directed Hypergraph Approach. *Proc. of the 50th Allerton Conference on Communications, Control, and Computing*, 2012.
12. M.R. Garey and D.S. Johnson. The Rectilinear Steiner Tree Problem is NP-Complete. *SIAM J. Appl. Math.*, 32:826–834, 1977.
13. R. Hassin, J. Monnot, and D. Segev. Approximation algorithms and hardness results for labeled connectivity problems. *J. Comb. Optim.*, 14:437–453, 2007.
14. A. Johansson and P. Dell'Acqua. Knowledge-based probability maps for covert pathfinding. In *MIG*, 339–350, 2010.
15. R.M. Karp. Reducibility among combinatorial problems. In R.E. Miller and J.W. Thatcher, Eds, *Complexity of Computer Computations*, 85–103. Plenum Press, NY, 1972.
16. P.N. Klein and R. Ravi. A nearly best-possible approximation algorithm for node-weighted steiner trees. *J. Algo.*, 19:104–115, 1995.
17. S.O. Krumke and H.-C. Wirth. On the minimum label spanning tree problem. *IPL*, 66:81–85, 1998.
18. B. Liu, O. Dousse, J. Wang, and A. Saipulla. Strong barrier coverage of wireless sensor networks. In *MobiHoc*, 411–420, 2008.
19. M. Marzouqi and R. Jarvis. New visibility-based path-planning approach for covert robotic navigation. *Robotica*, 24:759–773, 2006.
20. M. Marzouqi and R. Jarvis. Robotic covert path planning: A survey. In *RAM*, 77–82, 2011.
21. S. Meguerdichian, F. Koushanfar, M. Potkonjak, and M.B. Srivastava. Coverage problems in wireless ad-hoc sensor networks. In *INFOCOM*, 1380–1387, 2001.
22. J. Monnot. The labeled perfect matching in bipartite graphs. *IPL*, 96:81–88, 2005.
23. D. Peleg. Approximation algorithms for the label-cover$_{\max}$ and red-blue set cover problems. *J. Discrete Algo.*, 5:55–64, 2007.
24. N. Robertson and P.D. Seymour. Graph minors. ii. algorithmic aspects of tree-width. *J. Algo.*, 7:309–322, 1986.
25. S. Yuan, S. Varma, and J.P. Jue. Minimum-color path problems for reliability in mesh networks. In *INFOCOM*, 2658–2669, 2005.
26. P. Zhang, J.Y. Cai, L. Tang, and W. Zhao. Approximation and hardness results for label cut and related problems. *J. Comb. Optim.*, 21:192–208, 2011.

6 Figures

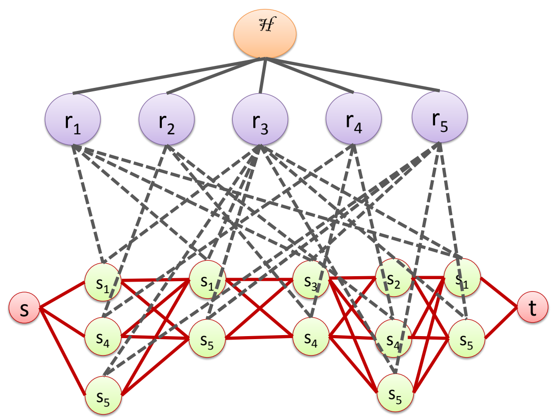

Fig. 1. Example of the gap-preserving reduction from `Red-Blue Set Cover` to `Secluded Path`. The `Red-Blue Set Cover` instance is given by $S_1 = \{b_1, b_2, b_3, r_1, r_3\}, S_2 = \{b_4, r_4\}, S_3 = \{b_3, r_1, r_5\}, S_4 = \{b_1, b_3, b_4, r_2, r_4\}, S_5 = \{b_1, b_2, b_4, b_5, r_3, r_5\}$. The optimal solution for `Red-Blue Set Cover` is $\{S_1, S_3, S_5\}$, which covers 3 red elements. The corresponding $s - t$ path is of cost $14 + 3 \cdot 5^3 = 389$. Note that an alternative solution of $\{S_1, S_4\}$ is of higher cost, covering 4 red elements, and the corresponding $s - t$ path is of cost $14 + 4 \cdot 5^3 = 514$.

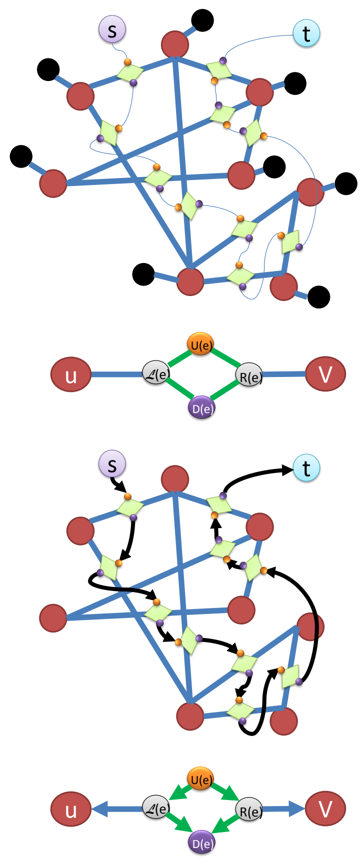

Fig. 2. Illustration of the reduction from `Vertex Cover`. (a) The weighted case. Top: the graph G with the diamond graph gadgets $g(e)$. The black nodes correspond to the heavy neighbor attached to each node $v \in V(G)$. Bottom: zoom into a single gadget. (b) The directed case. Directionality enforces visiting the gadgets in order, precluding the tour in $V(G)$ nodes.

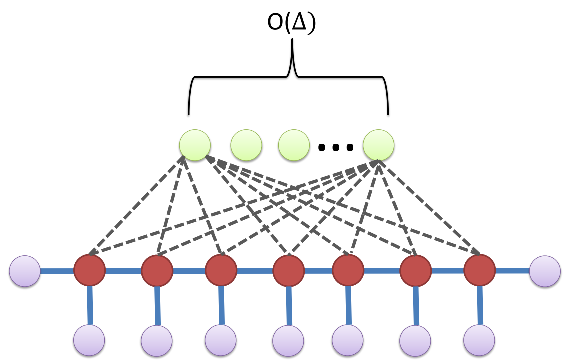

Fig. 3. Illustration of the gap between $q^*(\mathcal{S})$ and $\texttt{DegCost}^*(\mathcal{S})$. The light purple nodes correspond to the set of terminals $\mathcal{S}$.