



## Clearing directed subgraphs by mobile agents Variations on covering with paths



Dariusz Dereniowski<sup>a,\*</sup>, Andrzej Lingas<sup>b</sup>, Dorota Osula<sup>a</sup>, Mia Persson<sup>c</sup>,  
Paweł Żyliński<sup>d</sup>

<sup>a</sup> Faculty of Electronics, Telecommunications and Informatics, Gdańsk University of Technology, 80-233 Gdańsk, Poland

<sup>b</sup> Department of Computer Science, Lund University, 221 00 Lund, Sweden

<sup>c</sup> Department of Computer Science, Malmö University, 205 06 Malmö, Sweden

<sup>d</sup> Institute of Informatics, University of Gdańsk, 80-309 Gdańsk, Poland

### ARTICLE INFO

#### Article history:

Received 1 December 2017

Received in revised form 26 July 2018

Accepted 8 November 2018

Available online 22 November 2018

#### Keywords:

Covering with paths

FPT-algorithm

NP-hardness

Monomial

### ABSTRACT

We study several problems of clearing subgraphs by mobile agents in digraphs. The agents can move only along directed walks of a digraph and, depending on the variant, their initial positions may be pre-specified. In general, for a given subset  $S$  of vertices of a digraph  $D$  and a positive integer  $k$ , the objective is to determine whether there is a subgraph  $H = (\mathcal{V}_H, \mathcal{A}_H)$  of  $D$  such that (a)  $S \subseteq \mathcal{V}_H$ , (b)  $H$  is the union of  $k$  directed walks in  $D$ , and (c) the underlying graph of  $H$  includes a Steiner tree for  $S$  in  $D$ . Since a directed walk is a not necessarily a simple directed path, the problem is actually on covering with paths. We provide several results on the polynomial time tractability, hardness, and parameterized complexity of the problem. Our main fixed-parameter algorithm is randomized.

© 2018 Elsevier Inc. All rights reserved.

## 1. Introduction

Consider a city, after a snowstorm, where all streets have been buried in snow completely, leaving a number of facilities disconnected. For snow teams, distributed within the city, the main battle is usually first to re-establish connectedness between these facilities. This motivates us to introduce a number of (theoretical) *snow team* problems in graphs. Herein, in the introduction section, we shall formalize only one of them, leaving the other variants to be stated and discussed subsequently.

Let  $D = (\mathcal{V}, \mathcal{A}, \mathcal{F}, B)$  be a vertex-weighted digraph of order  $n$  and size  $m$ , with a subset  $\mathcal{F}$  of  $\mathcal{V}$  and a vertex-weight function  $B: \mathcal{V} \rightarrow \mathbb{N}$ , such that its underlying graph is connected. (Recall that the *underlying graph* of  $D$  is a simple graph with the same vertex set and its two vertices  $u$  and  $v$  being adjacent if and only if there is an arc between  $u$  and  $v$  in  $D$ .) The snow problem is modeled by  $D$  as follows. The vertices of  $D$  correspond to street crossings while its arcs correspond to (one-way) streets, the set  $\mathcal{F}$  corresponds to locations of facilities, and the set  $\mathcal{B} = B^{-1}(\mathbb{N}^+)$  corresponds to vertices, called from now on *snow team bases*, where a (positive) number of snow ploughs is placed (so we shall refer to the function  $B$  as a *plough-quantity* function). Let  $\mathbf{k}_B = \sum_{v \in \mathcal{V}} B(v)$  be the total number of snow ploughs placed in the digraph.

\* Corresponding author.

E-mail address: [deren@eti.pg.edu.pl](mailto:deren@eti.pg.edu.pl) (D. Dereniowski).

### The Snow Team problem (ST)

Do there exist  $\mathbf{k}_B$  directed walks in  $D$ , with exactly  $B(v)$  starting points at each vertex  $v \in \mathcal{V}$ , whose edges induce a subgraph  $H$  of  $D$  such that all vertices in  $\mathcal{F}$  belong to one connected component of the underlying graph of  $H$ ? Note that the  $\mathbf{k}_B$  directed walks may overlap in vertices and even edges.

The ST problem may be understood as a question, whether for  $\mathbf{k}_B$  snow ploughs, initially located at snow team bases in  $\mathcal{B} = B^{-1}(\mathbb{N}^+)$ , where the number of snow ploughs located at  $v \in \mathcal{B}$  is equal to  $B(v)$ , it is possible to follow  $\mathbf{k}_B$  walks in  $D$  clearing their arcs so that the underlying graph of the union of cleared walks includes a Steiner tree for all facilities in  $\mathcal{F}$ .

*Related work* The Snow Team problem is related to the problems of clearing connections by mobile agents placed at some vertices in a digraph, introduced by Levkopoulos et al. in [22]. In particular, the ST problem is a generalized variant of the Agent Clearing Tree (ACT) problem where one wants to determine a placement of the minimum number of mobile agents in a digraph  $D$  such that agents, allowed to move only along directed walks, can simultaneously clear some subgraph of  $D$  whose underlying graph includes a spanning tree of the underlying graph of  $D$ . In [22], the authors provided a simple 2-approximation algorithm for solving the Agent Clearing Tree problem, leaving its complexity status open.

All the aforementioned clearing problems are variants of the path cover problem in digraphs, where the objective is to find a minimum number of directed walks that cover all vertices (or edges) of a given digraph. Without any additional constraints, the problem was shown to be polynomially tractable by Ntafos and Hakimi in [26]. Several other variants involve additional constraints on walks as the part of the input, see [2,13,17,20,25–27] to mention just a few, some of them combined with relaxing the condition that all vertices of the digraph have to be covered by walks. In particular, we may be interested in covering only a given set of walks that themselves should appear as subwalks of some covering walks (polynomially tractable [26]). We may also be interested in covering only a given set of vertex pairs, where both elements of a pair should appear in the same order and in the same path in a solution (NP-complete even in acyclic digraphs [27]). Finally, for a given family  $\mathcal{S}$  of vertex subsets of  $D$ , we may be interested in covering only a representative from each of the subsets (NP-complete [27]).

A wider perspective locates our snow team problems as variants of the directed Steiner tree problem, where for a given edge-weighted directed graph  $D = (\mathcal{V}, \mathcal{A})$ , a root  $r \in \mathcal{V}$  and a set of facilities  $\mathcal{X} \subseteq \mathcal{V}$ , the objective is to find a minimum cost arborescence rooted at  $r$  and spanning all facilities in  $\mathcal{X}$  (equivalently, there exists a directed path from  $r$  to each facility in  $\mathcal{X}$ ) [8,34]. For some recent works and results related to this problem, see e.g. [1,12,16]. We also point out to a generalization of the Steiner tree problem in which pairs of facilities are given as an input and the goal is to find a minimum cost subgraph which provides a connection for each pair [8,10]. For some other generalizations, see e.g. [9,21,29–31]. Finally, we also remark a different cleaning problem introduced in [23] and related to the variants we study: cleaning a graph with brushes – for some recent works, see e.g. [6,7,15,24].

*Our results* We show that the Snow Team problem as well as some of its variants are fixed-parameter tractable. In particular, we prove that the ST problem admits a fixed-parameter randomized algorithm with respect to the total number  $l$  of facilities and snow team bases, running in  $2^{O(l)} \cdot \text{poly}(n)$  time, where  $\text{poly}(n)$  is a polynomial in the order  $n$  of the input graph (Section 2). The proof relies on the algebraic framework introduced by Koutis in [18]. On the other hand, we show that the ST problem (as well as some of its variants) is NP-complete, by a reduction from the Set Cover problem [14] (Section 3). Our result on NP-completeness of the ST problem implies NP-completeness of the Agent Clearing Tree problem studied in [22], where the complexity status of the latter has been posed as an open problem.

*Remark* Note that a weaker version of the ST problem with the connectivity requirement removed, that is, when we require each facility only to be connected to some snow team base, admits a polynomial-time solution by a straightforward reduction to the minimum path cover problem in directed graphs [26].

*Notation* A source of a directed graph is a vertex of indegree zero. The set of all source vertices in a directed graph  $D$  is denoted by  $s(D)$ . For a directed walk  $\pi$  in  $D$ , the set of vertices (arcs) of  $\pi$  is denoted by  $V(\pi)$  (resp.  $A(\pi)$ ), and its length – by  $|\pi|$ . For two directed walks  $\pi_1$  and  $\pi_2$  in  $D$ , where  $\pi_2$  starts at the ending point of  $\pi_1$ , the concatenation of  $\pi_1$  and  $\pi_2$  is denoted by  $\pi_1 \circ \pi_2$ .

Observe that in a border case, all non-zero length walks of snow ploughs start at the same vertex of the input digraph  $D = (\mathcal{V}, \mathcal{A}, \mathcal{F}, B)$ . Therefore, we may assume that the number of snow ploughs at any vertex is at most  $n - 1$ , that is,  $B(v) \leq n - 1$  for any  $v \in \mathcal{V}$ , and so the description of any input requires  $O(n \log n + m)$  space (recall  $m \geq n - 1$ ).

## 2. The ST problem is fixed-parameter tractable

In this section, we prove that the Snow Team problem is fixed-parameter tractable with respect to the number of facilities and snow team bases. The proof relies on the key fact (see Lemmas 2.1 and 2.2 below) that a restricted variant of the ST problem with the input  $D$  can be reduced to the detection of a particular directed subtree of ‘small’ order in the transitive closure  $\text{TC}(D)$  of  $D$ . By the *transitive closure* of a (simple) directed graph, we mean a (simple) directed graph on

the same vertices such that for each pair  $(u, v)$  of distinct vertices, there is an arc from  $u$  to  $v$  if and only if there is a directed path from  $u$  to  $v$  in the original graph.

We solve the latter tree detection problem by a reduction to the problem of testing whether some properly defined multivariate polynomial has a monomial with specific properties, essentially modifying the construction in [19] designed for undirected trees/graphs.

Let us consider the variant of the ST problem, which we shall refer to as the All-ST problem, where we restrict the input only to digraphs  $D = (\mathcal{V}, \mathcal{A}, \mathcal{F}, B)$  that satisfy  $\mathcal{B} = B^{-1}(\mathbb{N}^+) \subseteq \mathcal{F}$ . (In other words, snow team bases can be located only at some facilities.) Observe that  $D$  admits a positive answer to the ST problem if and only if there exists a subset  $\mathcal{B}'$  of  $\mathcal{B} \setminus \mathcal{F}$  such that the digraph  $D' = (\mathcal{V}, \mathcal{A}, \mathcal{F}', B')$ , where  $v \in \mathcal{F}'$  for  $v \in \mathcal{F} \cup \mathcal{B}'$  and  $v \notin \mathcal{F}'$  otherwise, and  $B'(v) = B(v)$  for  $v \in \mathcal{B}' \cup (\mathcal{F} \cap \mathcal{B})$  and  $B'(v) = 0$  otherwise, admits a positive answer to the All-ST problem. Therefore, we can immediately conclude with the following lemma.

**Lemma 2.1.** *Suppose that the All-ST problem can be solved in  $2^{O(k)} \cdot \text{poly}(n)$  time, where  $k$  is the number of facilities in the input (restricted) digraph of order  $n$ . Then, the ST problem can be solved in  $2^{O(l)} \cdot \text{poly}(n)$  time, where  $l$  is the total number of facilities and snow team bases in the input digraph of order  $n$ .  $\square$*

Taking into account the above lemma, we now focus on constructing an efficient fixed-parameter algorithm for the All-ST problem, with the *restricted* input digraph  $D = (\mathcal{V}, \mathcal{A}, \mathcal{F}, B)$  satisfying  $\mathcal{B} = B^{-1}(\mathbb{N}^+) \subseteq \mathcal{F}$ . Let  $\mathcal{W}$  be a set of walks (if any) that constitute a positive answer to the All-ST problem in  $D$ . We say that  $\mathcal{W}$  is *tree-like* if all walks in  $\mathcal{W}$  are strongly arc-distinct, that is, they are arc-distinct and if there is a walk in  $\mathcal{W}$  traversing the arc  $(u, v) \in \mathcal{A}$ , then there is no walk in  $\mathcal{W}$  traversing its complement  $(v, u) \in \mathcal{A}$ , and the underlying graph of their union is acyclic and includes a Steiner tree for  $\mathcal{F}$ . Notice that if  $\mathcal{W}$  is tree-like, then all walks in  $\mathcal{W}$  are just (simple) paths.

**Lemma 2.2.** *A (restricted) instance  $D = (\mathcal{V}, \mathcal{A}, \mathcal{F}, B)$  admits a positive answer to the All-ST problem if and only if the transitive closure  $\text{TC}(D) = (\mathcal{V}, \mathcal{A}', \mathcal{F}, B)$  of  $D$ , with the same subset  $\mathcal{F}$  and vertex-weight function  $B$ , admits a positive answer to the All-ST problem with a tree-like set of walks whose underlying graph is of order at most  $2|\mathcal{F}| - 1$ .*

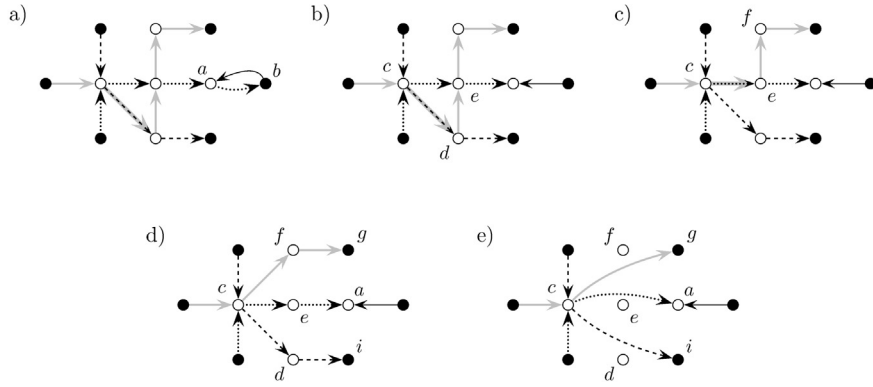
Since the transitive closure  $\text{TC}(D) = (\mathcal{V}, \mathcal{A}', \mathcal{F}, B)$  inherits the subset  $\mathcal{F}$  and the function  $B$  from the restricted instance  $D$ , we emphasize that  $\text{TC}(D)$  is a proper (restricted) instance to the All-ST problem.

**Proof.** ( $\Leftarrow$ ) It follows from the fact that a directed walk in the transitive closure  $\text{TC}(D)$  corresponds to a directed walk in  $D$ .

( $\Rightarrow$ ) Assume that the snow ploughs initially located at vertices in  $\mathcal{B}$ , with respect to the plough-quantity function  $B$ , can simultaneously follow  $\mathbf{k}_B$  directed walks  $\pi_1, \dots, \pi_{\mathbf{k}_B}$  whose edges induce a subgraph  $H$  of  $D$  such that the underlying graph of  $H$  includes a Steiner tree of  $\mathcal{F}$ . Consider now the same walks in the transitive closure  $\text{TC}(D)$ . To prove the existence of a tree-like solution of ‘small’ order, the idea is to transform these  $\mathbf{k}_B$  walks (if ever needed) into another strongly arc-distinct  $\mathbf{k}_B$  walks. The latter walks have the same starting points as the original ones (thus preserving the plough-quantity function  $B$ ) and the underlying graph of their union is a Steiner tree of  $\mathcal{F}$  (in the underlying graph of  $\text{TC}(D)$ ) having at most  $|\mathcal{F}| - 1$  non-terminal vertices.

Our transforming process is based on the following 2-step modification. First, assume without loss of generality that the walk  $\pi_1 = (v_1, \dots, v_{|\pi_1|+1})$  has an arc  $(v_t, v_{t+1})$  corresponding to an edge in the underlying graph  $H$  of  $\bigcup_{i=1}^{\mathbf{k}_B} \pi_i$  that belongs to a cycle (in  $H$ ), or both  $(v_t, v_{t+1})$  and its complement  $(v_{t+1}, v_t)$  are traversed by  $\pi_1$ , or there is another walk traversing  $(v_t, v_{t+1})$  or  $(v_{t+1}, v_t)$ , see Fig. 1 for an illustration. If  $t = |\pi_1|$ , then we shorten  $\pi_1$  by deleting its last arc  $(v_t, v_{t+1})$ . Otherwise, if  $t < |\pi_1|$ , then we replace the arcs  $(v_t, v_{t+1})$  and  $(v_{t+1}, v_{t+2})$  in  $\pi_1$  with the arc  $(v_t, v_{t+2})$  that exists because of the transitive closure. One can observe that the underlying graph of the new set of walks is connected, includes a Steiner tree of  $\mathcal{F}$ , and the vertex  $v_1$  remains the starting vertex of (the new)  $\pi_1$ . But, making a walk cycle-free or strongly arc-distinct may introduce another cycle in the underlying graph, or another multiply traversed arc, or another arc  $a$  such that both  $a$  and its complement are traversed. However, the length of the modified walk always decreases by one. Consequently, since the initial walks are of the finite lengths, we conclude that applying the above procedure multiple times eventually results in a tree-like set  $\Pi = \{\pi_1, \dots, \pi_{\mathbf{k}_B}\}$  of walks, being (simple) paths.

Assume now that in this set  $\Pi$  of strongly arc disjoint paths, there is a non-terminal vertex  $v$  of degree at most two in the underlying graph  $H$  of  $\bigcup_{i=1}^{\mathbf{k}_B} \pi_i$ . Without loss of generality assume that  $v$  belongs to the path  $\pi_1$ . Similarly as above, if  $\deg_H(v) = 1$ , then we shorten  $\pi_1$  by deleting its last arc. Otherwise, if  $\deg_H(v) = 2$  and  $v$  is not the endpoint of  $\pi_1$ , then modify  $\pi_1$  by replacing  $v$  together with the two arcs of  $\pi_1$  incident to it by the arc connecting the predecessor and successor of  $v$  in  $\pi_1$ , respectively. Observe that since  $v$  was a non-terminal vertex in the underlying graph, the underlying graph of (the new)  $\bigcup_{i=1}^{\mathbf{k}_B} \pi_i$  is another Steiner tree of  $\mathcal{F}$  (and does not include  $v$ ). Moreover, the above modification keeps paths strongly arc-distinct and does not change the starting vertex of  $\pi_1$ . Therefore, by subsequently replacing all such non-terminal vertices of degree at most two, we obtain a tree-like set of  $\mathbf{k}_B$  paths in the transitive closure  $\text{TC}(D)$  such that the underlying graph of their union is a Steiner tree of  $\mathcal{F}$  with no degree two vertices except those either belonging to  $\mathcal{F}$  or being end-vertices of exactly two paths (in  $\text{TC}(D)$ ). Therefore, we conclude that the number of non-terminal vertices in this underlying graph is at most  $|\mathcal{F}| - 1$ , which completes our proof of the lemma.



**Fig. 1.** Transforming walks into a tree-like set of walks. In the first step (a–d), applied three times, we first delete the arc  $(a, b)$ , then replace two arcs  $(c, d)$  and  $(d, e)$  with the arc  $(c, e)$ , and then two arcs  $(c, e)$  and  $(e, f)$  with the arc  $(c, f)$ . In the second step (d–e), we replace two arcs  $(c, d)$  and  $(d, i)$ , two arcs  $(c, e)$  and  $(e, a)$ , and two arcs  $(c, f)$  and  $(f, g)$ , with the arcs  $(c, i)$ ,  $(c, a)$ , and  $(c, g)$ , respectively.

Indeed, the bound is obvious if  $|\mathcal{F}| \leq 2$ . So assume now that  $|\mathcal{F}| \geq 3$ , the statement is valid for any set  $\mathcal{F}' \subset \mathcal{V}$  with  $0 \leq |\mathcal{F}'| < |\mathcal{F}|$ , and let  $\Pi$  be a tree-like set of paths in the transitive closure  $\text{TC}(D)$  such that the underlying graph  $H$  of their union is a Steiner tree  $T$  of  $\mathcal{F}$  with no degree two vertices except those either belonging to  $\mathcal{F}$  or being end-vertices of exactly two paths. Let  $v$  be a non-terminal vertex in  $H$  (if no such  $v$  exists, then there is nothing to prove), and let  $\Pi' \subseteq \Pi$  be the set of paths ending at vertex  $v$ . By deleting all path arcs with endpoint  $v$  for paths in  $\Pi'$  and by replacing two consecutive path arcs incident to  $v$  by the relevant arc connecting the predecessor and successor of  $v$  in  $\pi$ , respectively, for any path  $\pi \in \Pi \setminus \Pi'$ , we obtain the set  $\Pi'$  of strongly arc-distinct paths and a non-trivial partition  $\mathcal{F}_1 \cup \dots \cup \mathcal{F}_r = \mathcal{F}$  such that the underlying graph of their union consists of  $r \geq 2$  Steiner trees  $T_1, \dots, T_r$  of  $\mathcal{F}_1, \dots, \mathcal{F}_r$ , respectively, all of them with no degree two vertices except those either belonging to  $\mathcal{F}$  or being end-vertices of exactly two paths. By induction assumption, each tree  $T_i$  has at most  $|\mathcal{F}_i| - 1$  non-terminal vertices,  $i = 1, \dots, r$ , and so  $T$  has at most  $1 + \sum_{i=1}^r (|\mathcal{F}_i| - 1) \leq |\mathcal{F}| - 1$  non-terminal vertices since  $r \geq 2$ .  $\square$

Taking into account the above lemma, a given (restricted) instance  $D = (\mathcal{V}, \mathcal{A}, \mathcal{F}, B)$  of the All-ST problem can be transformed (in polynomial time) into the answer-equivalent (restricted) instance  $\text{TC}(D) = (\mathcal{V}, \mathcal{A}', \mathcal{F}, B)$  of the *tree-like-restricted* variant of the All-ST problem in which only tree-like plough paths that together visit at most  $2|\mathcal{F}| - 1$  vertices are allowed. Observe that  $\text{TC}(D) = (\mathcal{V}, \mathcal{A}', \mathcal{F}, B)$  admits a positive answer to the tree-like-restricted All-ST problem if and only if  $\text{TC}(D)$  has a subtree  $T = (\mathcal{V}_T, \mathcal{A}_T)$  of order at most  $2|\mathcal{F}| - 1$  and such that  $\mathcal{F} \subseteq \mathcal{V}_T$  and all edges of  $T$  can be traversed by at most  $\mathbf{k}_B$  snow ploughs following arc-distinct paths starting at vertices in  $\mathcal{B}$  (obeying the plough-quantity function  $B$ ). This motivates us to consider the following problem.

Let  $D = (\mathcal{V}, \mathcal{A}, \mathcal{F}, B)$  be a directed graph of order  $n$  and size  $m$ , with the subset  $\mathcal{F}$  of  $\mathcal{V}$  and a vertex-weight function  $B: \mathcal{V} \rightarrow \mathbb{N}$  such that  $B^{-1}(\mathbb{N}^+) \subseteq \mathcal{F}$ , and let  $T = (V, A, L)$  be a directed vertex-weighted tree of order  $\eta$ , with a vertex-weight function  $L: V \rightarrow \mathbb{N}$ .

### The Tree Pattern Embedding problem (TPE)

Does  $D$  have a subgraph  $H = (\mathcal{V}_H, \mathcal{A}_H)$  isomorphic to  $T$  such that  $\mathcal{F} \subseteq \mathcal{V}_H$  and  $L(v) \leq B(h(v))$  for any vertex  $v$  of  $T$ , where  $h$  is an isomorphism of  $T$  and  $H$ ?

In Subsection 2.1, we prove Theorem 2.1 given below which states that there is a randomized algorithm that solves the TPE problem in  $O^*(2^\eta)$  time, where the notation  $O^*$  suppresses polynomial terms in the order  $n$  of the input graph  $D$ . We point out that if the order  $\eta$  of  $T$  is less than  $|\mathcal{F}|$  or at least  $n + 1$ , then the problem becomes trivial, and so, in the following, we assume  $|\mathcal{F}| \leq \eta \leq n$ .

**Theorem 2.1.** *There is a randomized algorithm that solves the TPE problem with high probability in  $O^*(2^\eta)$  time.*

Suppose that for each vertex  $v \in V$ , the value  $L(v)$  corresponds to the number of snow ploughs located at  $v$  that are required to simultaneously traverse (clear) all arcs of  $T$ , in an arc-distinct manner, and  $T$  admits a positive answer to the TPE problem in the transitive closure  $\text{TC}(D) = (\mathcal{V}, \mathcal{A}', \mathcal{F}, B)$ . Then  $\text{TC}(D)$  admits a positive answer to the tree-like-restricted All-ST problem, which immediately implies that  $D$  admits a positive answer to the All-ST problem (by Lemma 2.2). Therefore, taking into account Theorem 2.1, we are now ready to present the main theorem of this section. For simplicity of presentation, we now assume that a (restricted) directed graph  $D = (\mathcal{V}, \mathcal{A}, \mathcal{F}, B)$  itself (not its transitive closure) is an instance of the tree-like-restricted All-ST problem.

**Theorem 2.2.** *There is a randomized algorithm that solves the tree-like-restricted All-ST problem for  $D = (\mathcal{V}, \mathcal{A}, \mathcal{F}, B)$  with high probability in  $O^*(144^{|\mathcal{F}|})$  time.*

**Proof.** Keeping in mind Lemma 2.2, we enumerate all undirected trees of order  $\eta$ , where  $|\mathcal{F}| \leq \eta \leq 2|\mathcal{F}| - 1$  (and  $\eta \leq n$ ); there are  $O(9^{|\mathcal{F}|})$  such candidates [28]. For each such a  $\eta$ -vertex candidate tree, we enumerate all orientations of its edges, in order to obtain a directed tree; there are  $2^{\eta-1}$  such orientations. Therefore, we have  $O(36^{|\mathcal{F}|})$  candidates for a directed oriented tree  $T$  of order  $\eta$ , where  $|\mathcal{F}| \leq \eta \leq 2|\mathcal{F}| - 1$ .

For each candidate  $T = (V, A)$ , we determine in  $O(\eta)$  time how many (at least) snow ploughs, together with their explicit locations at vertices in  $V$ , are needed to traverse all arcs of  $T$ , in an arc-distinct manner. This problem can be solved in linear time just by noting that the number of snow ploughs needed at a vertex  $v$  is equal to  $\max\{0, \deg_{\text{out}}(v) - \deg_{\text{in}}(v)\}$  (since arcs must be traversed in an arc-distinct manner). The locations of snow ploughs define a vertex-weight function  $L: V \rightarrow \mathbb{N}$ . We then solve the TPE problem with the instance  $D$  and  $T = (V, A, L)$  in  $O^*(2^\eta)$  time by Theorem 2.1.

As already observed, if  $T$  admits a positive answer to the TPE problem for  $D$ , then  $D$  admits a positive answer to the tree-like-restricted All-ST problem. Therefore, by deciding the TPE problem for each of  $O(36^{|\mathcal{F}|})$  candidates, taking into account the independence of any two tests, we obtain a randomized algorithm for the restricted ST problem with a running time  $O^*(144^{|\mathcal{F}|})$ .  $\square$

Taking into account Lemma 2.1, we immediately obtain the following corollary.

**Corollary 2.1.** *The ST problem admits a fixed-parameter randomized algorithm with respect to the total number  $l$  of facilities and snow team bases, running in  $2^{O(l)} \cdot \text{poly}(n)$  time, where  $n$  is the order of the input graph.*  $\square$

**Minimizing the number of used snow ploughs** The first natural variation on the Snow Team problem is its minimization variant, which we shall refer to as the min-ST problem, where for a given input  $n$ -vertex digraph  $D = (\mathcal{V}, \mathcal{A}, \mathcal{F}, B)$ , we wish to determine the minimum number of snow ploughs among those available at snow team bases in  $\mathcal{B} = B^{-1}(\mathbb{N}^+)$  that are enough to guarantee a positive answer to the (original) Snow Team problem in  $D$ . We claim that this problem also admits a fixed-parameter algorithm with respect to the total number  $l$  of facilities and snow bases, running in time  $2^{O(l)} \text{poly}(n)$ , and the solution is concealed in our algorithm for the ST problem. Namely, observe that by enumerating all directed trees of order at most  $|\mathcal{F}|$ , see the proof of Theorem 2.2, together with the relevant function  $L$ , and checking their embeddability in  $D$ , we accidentally solve this minimization problem: the embeddable tree with the minimum sum  $\sum_{v \in V} L(v)$  constitutes the answer to min-ST problem.

**Corollary 2.2.** *The min-ST problem admits a randomized fixed-parameter algorithm with respect to the total number  $l$  of facilities and snow team bases, running in  $2^{O(l)} \cdot \text{poly}(n)$  time, where  $n$  is the order of the input graph.*  $\square$

**Maximizing the number of re-connected facilities** In the case when for the input digraph  $D = (\mathcal{V}, \mathcal{A}, \mathcal{F}, B)$ , not all facilities can be re-connected into one component, that is,  $D$  admits a negative answer to the Snow Team problem, one can ask about the maximum number of facilities in  $\mathcal{F}$  that can be re-connected by snow ploughs located with respect to the plough-quantity function  $B$  [33]; we shall refer to this problem as the max-ST problem. Since we can enumerate all subsets of  $\mathcal{F}$  in  $O^*(2^{|\mathcal{F}|})$  time, taking into account Theorem 2.2, we obtain the following corollary.

**Corollary 2.3.** *The max-ST problem admits a randomized fixed-parameter algorithm with respect to the total number  $l$  of facilities and snow team bases, running in  $2^{O(l)} \cdot \text{poly}(n)$  time, where  $n$  is the order of the input graph.*  $\square$

**No pre-specified positions of snow ploughs** Finally, another natural variant of the Snow Team problem is to allow any snow plough to start at any vertex. Formally, we define the following problem.

### The Snow Team problem with Unspecified snow team bases (STU)

Given a subset  $\mathcal{F}$  of  $\mathcal{V}$  and an integer  $k \geq 1$ , do there exist  $k$  directed walks in a digraph  $D = (\mathcal{V}, \mathcal{A})$  whose edges induce a subgraph  $H$  of  $D$  such that the set  $\mathcal{F}$  is a subset of the vertex set of  $H$  and the underlying graph of  $H$  is connected? Again, let us emphasize that these  $k$  directed walks may overlap in vertices or edges.

We claim that for the STU problem, there is also a randomized algorithm with the running time  $2^{O(k+l)} \cdot \text{poly}(n)$ , where  $l = |\mathcal{F}|$  is the number of facilities, and  $n$  is the order of the input graph. The solution is analogous to that for the ST problem. Namely, one can prove a counterpart of Lemma 2.2 which allows us to restrict ourselves to the restricted variant where only order  $O(k+l)$  tree-like solutions are allowed. Then, the restricted variant is solved also using the algorithm for the TPE problem as a subroutine: the function  $B$  is the constant function  $B(v) = n$ , and among all directed tree candidates, we check only those with  $\sum_{v \in V} L(v) \leq k$ . We omit details.





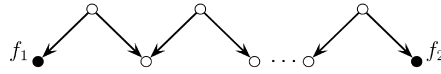


Fig. 2. Two facilities  $f_1$  and  $f_2$  require  $n - 1$  snow ploughs, where  $n$  is the order of the digraph.

**Corollary 2.4.** *The STU problem admits a randomized fixed-parameter algorithm with respect to the number  $l$  of facilities and the number  $k$  of snow ploughs, running in  $2^{O(k+l)} \cdot \text{poly}(n)$  time, where  $n$  is the order of the input graph.  $\square$*

Observe that if the number  $k$  of available snow ploughs is not a part of the input, that is, we ask about the minimum number of walks whose underlying graph includes a Steiner tree for the set of facilities, then this problem does not seem to be fixed-parameter tractable with respect to only the number of facilities. This follows from the fact that the minimum number of snow ploughs is unrelated to the number of facilities in the sense that even for two facilities to be connected, a lot of snow ploughs may be required, see Fig. 2 for an illustration. However, a weakness of this example is its (weak) connectivity, and so, without any strong evidence, we conjecture that if restricted only to strongly connected digraphs, the aforementioned problem becomes then fixed-parameter tractable with respect to only the number of facilities.

### 2.1. The Tree Pattern Embedding problem

In this section, we solve the TPE problem by providing a randomized polynomial-time algorithm when the parameter  $\eta$  is fixed. Our algorithm is based on the recent algebraic technique using the concepts of monotone arithmetic circuits and monomials, introduced by Koutis in [18], developed by Williams and Koutis in [19,32], and adapted to some other graph problems, e.g., [3–5,11].

A (monotone) arithmetic circuit is a directed acyclic graph where each leaf (i.e., vertex of in-degree 0) is labeled either with a variable or a real non-negative constant (input gates), each non-leaf vertex is labeled either with  $+$  (an addition gate with an unbounded fan-in) or with  $\times$  (a multiplication gate with fan-in two), and where a single vertex is distinguished (the output gate). Each vertex (gate) of the circuit represents (computes) a polynomial – these are naturally defined by induction on the structure of the circuit starting from its input gates – and we say that a polynomial is represented (computed) by an arithmetic circuit if it is represented (computed) by the output gate of the circuit. Finally, a polynomial that is just a product of variables is called a monomial, and a monomial in which each variable occurs at most once is termed a multilinear monomial [18,32].

We shall use a slight generalization of the main results of Koutis and Williams in [18,32], provided by them in Lemma 1 in [19], which, in terms of our notation, can be expressed as follows.

**Lemma 2.3.** ([19]) *Let  $P(x_1, \dots, x_n, z)$  be a polynomial represented by a monotone arithmetic circuit of size  $s(n)$  and let  $t$  be a non-negative integer. There is a randomized algorithm that for input  $P$  runs in  $O^*(2^k t^2 s(n))$  time and outputs “YES” with high probability if there is a monomial of the form  $z^t Q(x_1, \dots, x_n)$ , where  $Q(x_1, \dots, x_n)$  is a multilinear monomial of degree at most  $k$ , in the sum-product expansion of  $P$ , and always outputs “NO” if there is no such monomial  $z^t Q(x_1, \dots, x_n)$  in the expansion.  $\square$*

Taking into account the above lemma, for the input digraph  $D = (\mathcal{V}, \mathcal{A}, \mathcal{F}, B)$  and directed tree  $T = (V, A, L)$ , the idea is to construct an appropriate polynomial  $Q(X, z)$  such that  $Q(X, z)$  contains a monomial of the form  $z^{|\mathcal{S}|} b(X)$ , where  $b(X)$  is a multilinear polynomial with exactly  $|\mathcal{V}|$  variables in  $X$  and  $\mathcal{S} = \mathcal{F} \cup B^{-1}(\mathbb{N}^+)$ , if and only if the TPE problem has a solution for the input  $D$  and  $T$  (see Lemma 2.4 below). Intuitively, the fact that  $b(X)$  is a multilinear polynomial ensures that no two vertices of  $T$  are mapped into the same vertex of  $D$  in the corresponding embedding. Moreover, the requirement that  $b(X)$  has exactly  $|\mathcal{V}|$  variables ensures that no vertex of  $T$  is missing in the mapping and no vertex is mapped twice. Finally, the variable  $z$  ‘counts’ the number of properly mapped vertices according to the functions  $L$  in  $T$  and  $B$  in  $D$ .

**Polynomial construction** Let  $D = (\mathcal{V}, \mathcal{A}, \mathcal{F}, B)$  be a directed graph, with a subset  $\mathcal{F}$  of  $\mathcal{V}$  and a vertex-weight function  $B: \mathcal{V} \rightarrow \mathbb{N}$ , and let  $T = (V, A, L)$  be a directed vertex-weighted tree of order  $\eta$ , with a vertex-weight function  $L: V \rightarrow \mathbb{N}$ . We consider  $T$  to be rooted at a vertex  $r \in V$ , and for a non-root vertex  $v$  of  $T$ , we denote the parent of  $v$  in  $T$  by  $p(v)$ . Now, for  $v \in V$ , define two sets  $N_T^+(v)$  and  $N_T^-(v)$ :

$$N_T^+(v) = \{u \in V \mid (u, v) \in A \text{ and } u \neq p(v)\},$$

$$N_T^-(v) = \{u \in V \mid (v, u) \in A \text{ and } u \neq p(v)\}.$$

The idea is to treat  $T$  as a ‘pattern’ that we try to embed into the digraph  $D$ , with respect to a subset  $\mathcal{F}$  and functions  $B, L$ . Denote  $\mathcal{S} = \mathcal{F} \cup B^{-1}(\mathbb{N}^+)$  for brevity. We say that  $T$  has an  $\mathcal{S}$ -embedding into  $D$  if the following holds (these are the formal conditions that need to be satisfied for the embedding to be correct):

- (E1) There exists an injective function (homomorphism)  $f: V \rightarrow \mathcal{V}$  such that if  $(u, v) \in A$ , then  $(f(u), f(v)) \in \mathcal{A}$ .
- (E2)  $\mathcal{S} \subseteq f(V)$ , where  $f(V) = \{f(v) \mid v \in V\}$ .
- (E3)  $L(v) \leq B(f(v))$  for any  $v \in V$ .

First, for  $\mathcal{S} \subseteq \mathcal{V}$ ,  $w \in \mathcal{V}$  and  $u \in V$ , we introduce a particular indicator function, used for fulfilling Conditions (E2) and (E3):

$$\mathbf{z}_{\mathcal{S}}(u, w) = \begin{cases} z, & \text{if } w \in \mathcal{S} \text{ and } L(u) \leq B(w), \\ 1, & \text{if } w \notin \mathcal{S} \text{ and } L(u) \leq B(w), \\ 0, & \text{otherwise, i.e., if } L(u) > B(w). \end{cases}$$

Next, following [19], we define a polynomial  $Q(X, T)$  that we then use to test existence of a desired  $\mathcal{S}$ -embedding of  $T$  in  $D$ . Namely, we root  $T$  at any vertex  $r \in V$ . Now, a polynomial  $Q_{u,w}(X)$ , for a subtree  $T_u$  of  $T$  rooted at  $u \in V$  and for a vertex  $w \in \mathcal{V}$ , is defined inductively (in a bottom up fashion on  $T$ ) as follows. For each  $u \in V$  and for each  $w \in \mathcal{V}$ : if  $u$  is a leaf in  $T$ , then

$$Q_{u,w}(X) = \mathbf{z}_{\mathcal{S}}(u, w) \cdot x_w, \tag{1}$$

and if  $u$  is not a leaf in  $T$ , then

$$Q_{u,w}(X) = \begin{cases} \mathbf{z}_{\mathcal{S}}(u, w) \cdot x_w \cdot Q_{u,w}^+(X) \cdot Q_{u,w}^-(X), & \text{if } N_T^-(u) \neq \emptyset \wedge N_T^+(u) \neq \emptyset, \\ \mathbf{z}_{\mathcal{S}}(u, w) \cdot x_w \cdot Q_{u,w}^+(X), & \text{if } N_T^-(u) = \emptyset, \\ \mathbf{z}_{\mathcal{S}}(u, w) \cdot x_w \cdot Q_{u,w}^-(X), & \text{if } N_T^+(u) = \emptyset, \end{cases} \tag{2}$$

where

$$Q_{u,w}^+(X) = \prod_{v \in N_T^+(u)} \left( \sum_{(w', w') \in \mathcal{A}} Q_{v,w'}(X) \right), \tag{3}$$

$$Q_{u,w}^-(X) = \prod_{v \in N_T^-(u)} \left( \sum_{(w, w') \in \mathcal{A}} Q_{v,w'}(X) \right). \tag{4}$$

Finally, the polynomial  $Q(X, z)$  is as follows:

$$Q(X, z) = \sum_{w \in \mathcal{V}} Q_{r,w}(X). \tag{5}$$

We have the following lemma.

**Lemma 2.4.** *The polynomial  $Q(X, z)$  contains a monomial of the form  $z^{|\mathcal{S}|} b(X)$ , where  $b(X)$  is a multilinear polynomial with exactly  $\eta$  variables in  $X$ , if and only if the  $\eta$ -vertex tree  $T$  has an  $\mathcal{S}$ -embedding into  $D$ .*

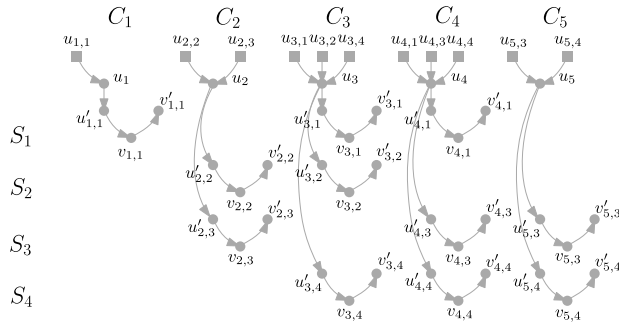
**Proof.** Consider a vertex  $u$  of  $T$  and assume that the subtree  $T_u$  is of order  $j$ . By a straightforward induction in the size of a subtree we state the following observation. A monomial, call it  $z^q x_{w_1} \cdots x_{w_j}$  for this subtree  $T_u$ , where  $w_i \in \mathcal{V}$  for each  $i \in \{1, \dots, j\}$ , is present in  $Q_{u,w_1}(X)$  if and only if the three following conditions hold.

- (i) There exists a homomorphism  $f_u$  from the vertices of  $T_u$  to  $w_1, \dots, w_j$  such that  $f_u(u) = w_1$ .
- (ii)  $|\mathcal{S} \cap \{w_1, \dots, w_j\}| \leq q$  and the equality holds if  $w_1, \dots, w_j$  are pairwise different.
- (iii)  $L(v) \leq B(f_u(v))$  for any vertex  $v$  of  $T_u$ .

The fact that  $f_u$  is a homomorphism follows from the observation that, during construction of  $Q_{u,w_1}(X)$  in (3) and (4), a neighbor  $v$  of  $u$  is mapped to a node  $w'$  of  $D$  in such a way that if  $(v, u) \in A$  then  $(w', w) \in \mathcal{A}$  (see (3)), and if  $(u, v) \in A$  then  $(w, w') \in \mathcal{A}$  (see (4)). Conditions (ii) and (iii) are ensured by appropriate usage of the indicator function in (1), namely, if  $u$  is mapped to  $w$  in a homomorphism corresponding to  $Q_{u,w}(X)$ , then we add the multiplicative factor of  $z$  to  $Q_{u,w}(X)$  provided that  $L(v) \leq B(w)$ .

Thus, we obtain that  $Q(X, z)$  has a multilinear polynomial  $z^{|\mathcal{S}|} x_{w_1} \cdots x_{w_\eta}$  if and only if  $T$  has an  $\mathcal{S}$ -embedding into  $D$ .  $\square$

Observe that the polynomial  $Q(X, z)$  and the auxiliary polynomials  $Q_{u,w}^+(X)$ ,  $Q_{u,w}^-(X)$  can be represented by a monotone arithmetic circuit of size polynomial in the order  $n$  of the input digraph  $D$ . To start with, we need  $n + 1$  input gates for the variables corresponding to vertices of  $D$ , and the auxiliary variable  $z$ . With each of the aforementioned polynomials, we associate a gate representing it, which gives in total  $O(\eta n)$  such gates. In order to implement the recurrences defining the polynomials, assuming unbounded fan-in of addition gates, we need  $O(n)$  auxiliary gates for each recurrence involving large products. Thus, the resulting circuit is of size  $O(n^3)$ . Hence, by Lemma 2.3 combined with Lemma 2.4, we conclude that the existence of an  $\mathcal{S}$ -embedding of the  $\eta$ -vertex tree  $T$  into  $D$  can be decided in  $O^*(2^\eta)$  time. Consequently, we obtain Theorem 2.1 by the definition of an  $\mathcal{S}$ -embedding.



**Fig. 3.** Step 1: construction of the element components. Here,  $\mathcal{U} = \{1, 2, 3, 4, 5\}$  and  $\mathcal{S} = \{S_1, S_2, S_3, S_4\}$ , where  $S_1 = \{1, 3, 4\}$ ,  $S_2 = \{2, 3\}$ ,  $S_3 = \{2, 4, 5\}$  and  $S_4 = \{3, 4, 5\}$ . The corresponding sets of indices are:  $I_1 = \{1\}$ ,  $I_2 = \{2, 3\}$ ,  $I_3 = \{1, 2, 4\}$ ,  $I_4 = \{1, 3, 4\}$  and  $I_5 = \{3, 4\}$ . For  $k=2$ , there exists a set cover  $\mathcal{X}^I = \{1, 3\}$  of size 2, consisting of the sets  $S_1$  and  $S_3$ .

*Remark* The above approach can be adapted to the case when we want to embed a directed forest  $T = (\mathcal{V}, \mathcal{A}, \mathcal{F}, B)$  of order  $\eta$  into a directed graph. All we need is to build a relevant polynomial for each rooted directed tree-component of  $T$ , and then to consider the product  $S(X, T)$  of these polynomials, asking about the existence of a monomial of the form  $z^{|\mathcal{S}|}b(X)$ , where  $b(X)$  is a multilinear polynomial with exactly  $\eta$  variables in  $X$ . Also, by a similar approach, we may consider and can solve (simpler) variants of our embedding problem without the weight function subset  $\mathcal{F}$  or without the weight functions  $B$  and  $L$ ; details are omitted.

### 3. The Snow Team problem is hard

In this section, we prove that the Snow Team problem is NP-complete by describing a polynomial-time reduction from the Set Cover problem.

Let  $\mathcal{U} = \{u_1, \dots, u_n\}$  be a set of  $n$  items and let  $\mathcal{S} = \{S_1, \dots, S_m\}$  be a family of  $m$  sets containing the items in  $\mathcal{U}$ , i.e., each  $S_t \subseteq \mathcal{U}$ , such that each element in  $\mathcal{U}$  belongs to at least one set from  $\mathcal{S}$ . A  $k$ -element subset of  $\mathcal{S}$ , whose union is equal to the whole universe  $\mathcal{U}$ , is called a *set cover of size k*.

#### The Set Cover problem (SC)

Given  $\mathcal{U}, \mathcal{S}$  and a positive integer  $k$ , does there exist a set cover of a size  $k$ ?

The Set Cover problem is well known to be NP-complete [14]. We are going to prove that for a given  $\mathcal{U} = \{u_1, \dots, u_n\}, \mathcal{S} = \{S_1, \dots, S_m\}$  and an integer  $k$ , there exists a set cover of size  $k$  if and only if there is a solution for the ST problem in the appropriately constructed acyclic digraph  $D_{SC} = (\mathcal{V}, \mathcal{A}, \mathcal{F}, B)$ . Basically, in this graph, for each element  $u \in \mathcal{U}$ , there is a gadget  $C_u$  being the union of the number of ‘vertical’ paths equal to the number of sets that  $u$  appears in. Also, there is one ‘spanning’ gadget including  $m$  ‘horizontal’ paths  $PH_t, t = 1, \dots, m$ , each of which visits the relevant gadget  $C_u$  if the element  $u$  belongs to the set  $S_t$ . In our construction, all vertices are facilities, and snow teams are located only at source vertices, one team at each source vertex (see details below).

In the following, we assume  $\mathcal{U} = \{1, \dots, n\}$  and that elements in  $S_t = \{x_1, \dots, x_{\ell(t)}\}$  are sorted in the ascending order, where  $\ell(t)$  denotes the size of  $S_t, t \in \{1, \dots, m\}$ . Also, we denote a solution to the set cover problem with the input  $(\mathcal{U}, \mathcal{S}, k)$  by  $\mathcal{X}^I$  which is encoded as a subset of  $\{1, \dots, m\}$ , where  $t \in \mathcal{X}^I$  if and only if  $S_t$  belongs to the  $k$ -element subset of  $\mathcal{S}$  forming the set cover. Finally, for simplicity of presentation, we denote a set  $\{1, \dots, l\}$  of indices by  $[l]$ .

*Digraph construction* For  $i \in \mathcal{U}$ , let  $I_i$  be the set of all indices of subsets from  $\mathcal{S}$  which contain  $i$ :  $I_i = \{j \mid i \in S_j\}$ . First, for every  $i \in \mathcal{U}$ , we introduce the following four sets of vertices (see Fig. 3 for an illustration):

- $U_i = \{u_i\} \cup \{u_{i,j} : j \in I_i\}$ ,
- $U'_i = \{u'_{i,j} : j \in I_i\}$ ,
- $V_i = \{v_{i,j} : j \in I_i\}$ ,
- $V'_i = \{v'_{i,j} : j \in I_i\}$ .

Next, we introduce an additional set  $Z = \{z, z_1, \dots, z_k\}$  of vertices corresponding to the input integer  $k$ . Finally, for every  $i \in \mathcal{U}$  and  $j \in I_i$ , we create a directed path  $PV_{i,j} = (u_{i,j}, u_i, u'_{i,j}, v_{i,j}, v'_{i,j})$ ; we refer to these paths as *vertical*. Observe that for  $i \in \mathcal{U}$ , the union  $\bigcup_{j \in I_i} V(PV_{i,j})$  induces the directed subgraph which we shall refer to as the *i-th element component*  $C_i$ . This finishes the first step of our construction.



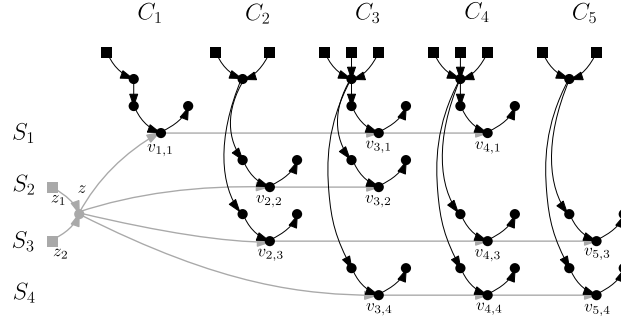


Fig. 4. (Cont. Fig. 3) Step 2: four horizontal paths connecting all components and two new source vertices  $z_1, z_2$  are added (recall  $k = 2$ ).

For the second step (see Fig. 4 for an illustration), for  $t \in [m]$ , we consider the set  $S_t = \{x_1, \dots, x_{\ell(t)}\}$  – recall that the elements in  $S_t$  are sorted in the ascending order – and construct a directed path  $PH_t = (z, v_{x_1,t}, v_{x_2,t}, \dots, v_{x_{\ell(t)},t})$ ; these paths are called *horizontal*. Next, we add  $k$  additional arcs, namely, for each  $l \in [k]$ , we add the arc  $(z_l, z)$ .

With our 2-step construction, we have built the directed graph  $D_{SC} = (\mathcal{V}, \mathcal{A})$ , where

$$\mathcal{V} = \bigcup_{i \in [n]} V(C_i) \cup Z \quad \text{and} \quad \mathcal{A} = \{(z_i, z) \mid i \in [k]\} \cup \bigcup_{i \in [m]} E(C_i) \cup \bigcup_{i \in [m]} E(PH_i).$$

We finalize our construction by defining the subset  $\mathcal{F}$  of  $\mathcal{V}$  and the function  $B: \mathcal{V} \rightarrow \mathbb{N}$ . Specifically, we set  $\mathcal{F} = \mathcal{V}$ , and  $B(v) = 1$  if and only if  $v \in \mathcal{V}$  is a source vertex in  $D_{SC}$ :

$$B(v) = \begin{cases} 1, & \text{if } v \in \{u_{i,j} \mid i \in \mathcal{U}, j \in I_i\} \cup \{z_1, \dots, z_k\} \\ 0, & \text{otherwise.} \end{cases}$$

Therefore,  $\mathcal{F} = \mathcal{V}$  and  $\mathcal{B} = B^{-1}(\mathbb{N}^+)$  is the set of all source vertices in  $D_{SC}$ . In particular, there is exactly one snow plough at each source vertex of  $D_{SC}$ , and so  $\mathbf{k}_B = \sum_{v \in \mathcal{V}} B(v) = |s(D_{SC})|$ , which equals to  $k + \sum_{i=1}^m |S_i|$  by the construction of  $D_{SC}$ .

Clearly, our reduction takes polynomial time. The order of  $D_{SC}$  is equal to  $4 \cdot \sum_{i=1}^m |S_i| + n + k + 1 = O(nm + k)$ , its size also equals  $O(nm + k)$ , and the descriptions of the set  $\mathcal{F}$  and the function  $B$  require  $O(nm + k)$  space either. Finally, observe that  $D_{SC}$  is acyclic and its underlying graph is connected. The latter observation follows from the fact that  $\mathcal{S}$  is a family of sets whose union is  $\mathcal{U}$ , and each element in  $\mathcal{U}$  belongs to at least one set from  $\mathcal{S}$ .

### 3.1. Direct implication

First, we are going to prove the direct implication.

**Lemma 3.1.** *Let  $(\mathcal{U}, \mathcal{S}, k)$  be an instance of the SC problem. If there exists a set cover of size  $k$  for  $\mathcal{U}$  and  $\mathcal{S}$ , then there exists a solution to the ST problem for the digraph  $D_{SC} = (\mathcal{V}, \mathcal{A}, \mathcal{F}, B)$ .*

**Proof.** (See Fig. 5 for an illustration.) Let  $\mathcal{X}^I = \{\xi(1), \dots, \xi(k)\}$  be a solution to the SC problem for  $(\mathcal{U}, \mathcal{S}, k)$ . We now construct a solution  $\mathcal{W}$  to the ST problem to consist of the following paths:

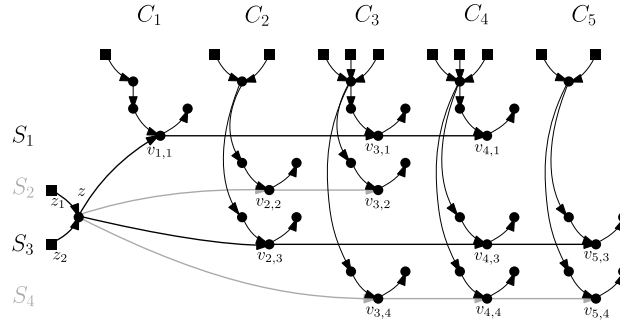
$$PV_{i,j} = (u_{i,j}, u_i, u'_{i,j}, v_{i,j}, v'_{i,j}), \quad \text{for } i \in \mathcal{U}, j \in I_i,$$

and

$$(z_t, z) \circ PH_{\xi(t)} = (z_t, z, v_{x_1, \xi(t)}, v_{x_2, \xi(t)}, \dots, v_{x_{\ell(t)}, \xi(t)}), \quad \text{for } t \in [k],$$

where  $x_1, \dots, x_{\ell(t)}$  are the (ordered) elements of the set  $S_t \in \mathcal{S}$ , with  $|S_t| = \ell(t)$ .

Clearly, by the construction, the set  $\mathcal{W}$  consists of  $\mathbf{k}_B = |s(D_{SC})|$  paths that together cover all vertices of  $D_{SC}$  and each of which starts at a distinct vertex in  $\mathcal{B} = B^{-1}(\mathbb{N}^+)$ . Thus, it remains to prove that these paths induce a subgraph  $H$  of  $D_{SC}$  whose underlying graph is connected. First, by the definition of the paths  $PV_{i,j}$ , observe that each element component induces a subgraph whose underlying graph is connected. Hence, it is enough to argue that for each element component  $C_i$ , where  $i \in \mathcal{U}$ , there exists a directed path in  $H$  that connects  $z$  with a vertex of  $C_i$ . Suppose for a contradiction that this is not the case for the  $i$ -th element component  $C_i$ , for some  $i \in \mathcal{U}$ . That means that no vertex  $v_{i,j}$ , for any  $j \in I_i$ , is lying on any of the chosen horizontal paths in  $H$ . But that means, by the choice of the horizontal paths and the fact that  $\mathcal{X}^I$  is a solution to the SC problem, that  $i$  does not belong to  $\bigcup_{j \in \mathcal{X}^I} S_j$ , a contradiction.  $\square$



**Fig. 5.** (Cont. Fig. 4) The fact that the set cover contains all elements in  $\mathcal{U}$  guarantees that the corresponding subgraph  $H$  of  $D$  is connected. For  $k = 2$ , there exists a set cover  $\mathcal{X}^l = \{1, 3\}$  of size 2, corresponding to set  $S_1$  and  $S_3$ , respectively.

### 3.2. Converse implication

In this subsection, to complete the NP-completeness proof of the ST problem, we are going to prove the converse implication, in a sequence of lemmas. Recall that in our digraph  $D_{SC} = (\mathcal{V}, \mathcal{A}, \mathcal{F}, B)$ , we have

$$s(D_{SC}) = \{z_1, \dots, z_k\} \cup \bigcup_{i \in \mathcal{U}} \{u_{i,j} \mid j \in I_i\},$$

and we set  $B(v) = 1$  for each  $v \in s(D_{SC})$  and  $B(v) = 0$  otherwise, and so  $\mathbf{k}_B = |s(D_{SC})|$ . Thus, keeping in mind that  $D_{SC}$  is acyclic, any solution  $\mathcal{W}$  to the ST problem for  $D_{SC}$  consists of  $\mathbf{k}_B$  paths that start at distinct source vertices in  $s(D_{SC})$ ; in the following,  $\pi(v) \in \mathcal{W}$  denotes the unique path of  $\mathcal{W}$  that starts at a source vertex  $v \in s(D_{SC})$ .

**Lemma 3.2.** *Suppose that the ST problem admits a positive answer for the input graph  $D_{SC}$ . Then, there exists a solution  $\{\pi(v) \mid v \in s(D_{SC})\}$  to the ST problem for  $D_{SC}$  such that for each  $i \in \mathcal{U}$  and  $j \in I_i$ ,  $V(\pi(u_{i,j})) \subseteq V(C_i)$ .*

**Proof.** Let  $\mathcal{W}$  be a solution to the ST problem for  $D_{SC}$  and assume that in  $\mathcal{W}$ , for some  $i \in \mathcal{U}$  and  $j \in I_i$ , we have  $\pi(u_{i,j}) = (u_{i,j}, u_i, u'_{i,j'}, v_{i,j'}, v_{i',j'}) \circ P$  for some (possibly empty) path  $P$ , that is, the arc  $(v_{i,j'}, v_{i',j'}) \in A(\pi(u_{i,j}))$  and so  $\pi(u_{i,j})$  'leaves'  $C_i$  at  $v_{i,j'}$  by visiting vertex  $v_{i',j'}$ , for some  $i' > i$ . We will argue that we may obtain another solution to the ST problem in which  $V(\pi(u_{i,j})) \subseteq V(C_i)$  as required by the lemma. The idea is to modify two paths in  $\mathcal{W}$  maintaining the following invariant: each path that is a subgraph of an element component remains a subgraph of this element component. Thus, since this modification can be repeated for any  $i \in \mathcal{U}$  and  $j \in I_i$ , this is sufficient to prove our claim.

Since  $\mathcal{F} = \mathcal{V}$ , we have  $v'_{i,j'} \in \mathcal{F}$  and hence there exists another path  $\pi(v) \in \mathcal{W}$  for some  $v \in s(D_{SC})$  such that  $\pi(v) = P' \circ (v_{i,j'}, v'_{i,j'})$ , for some path  $P'$  in  $D_{SC}$ . We then modify the set  $\mathcal{W}$  of paths by removing the two paths  $\pi(u_{i,j})$  and  $\pi(v)$ , and then adding the following two paths:  $(u_{i,j}, u_i, u'_{i,j'}, v_{i,j'}, v'_{i,j'})$  and  $P' \circ (v_{i,j'}, v'_{i,j'}) \circ P$ .

Observe that any two paths that start at vertices in  $U_i \setminus \{u_i\}$  have only the vertex  $u_i$  in common – this is due to the fact that the vertices in  $U'_i$  are only reachable with paths that start at vertices in  $U_i \setminus \{u_i\}$  and  $|U'_i| = |U_i \setminus \{u_i\}| = |B^{-1}(1) \cap V(C_i)|$ . Consequently, the vertex set  $V(\pi(v))$  of the original path  $\pi(v)$  is not a subset of a single element component, and hence, after the modification, each path whose vertex set is a subset of an element component keeps this property as required. Clearly, all vertices of  $D$  are covered in the new solution and, since the two new paths also share the vertex  $v_{i,j'}$ , the modified set of paths also induces a connected spanning subgraph of the underlying graph of  $D$ .  $\square$

**Lemma 3.3.** *Suppose that the ST problem admits a positive answer for the input graph  $D_{SC}$ . Then, there exists a solution  $\{\pi(v) \mid v \in s(D_{SC})\}$  to the ST problem for  $D_{SC}$  such that:*

- for each  $i \in \mathcal{U}$  and  $j \in I_i$ , we have  $\pi(u_{i,j}) = PV_{i,j}$ ;
- for each  $t \in [k]$ , we have  $\pi(z_t) = (z_t, z) \circ PH_l$  for some  $l \in [m]$ .

Before we proceed with the proof of Lemma 3.3, note that Property (a) implies that the paths  $\pi(u_{i,j})$ , where  $i \in \mathcal{U}$  and  $j \in I_i$ , visit all and only vertices of all components  $C_i$ ,  $i \in \mathcal{U}$ . Therefore, since these components have no vertices in common, the underlying graph becomes connected only thanks to the paths  $\pi(v)$ ,  $v \in Z \setminus \{z\}$ , which Property (b) refers to.

**Proof.** (a) Consider any solution, say  $\mathcal{W}$ , to the ST problem for  $D_{SC}$  and assume that in  $\mathcal{W}$ , for some  $i \in \mathcal{U}$  and  $j \in I_i$ , we have  $\pi(u_{i,j}) \neq PV_{i,j}$ ; we shall refer to  $\pi(u_{i,j})$  as well as to any other path  $\pi(u_{i',j'})$  such that  $\pi(u_{i',j'}) \neq PV_{i',j'}$  as *inconsistent*. By Lemma 3.2,  $\pi(u_{i,j}) \subseteq V(C_i)$  and hence there exists  $j' \in I_i \setminus \{j\}$  such that  $\pi(u_{i,j}) = (u_{i,j}, u_i, u'_{i,j'}, v_{i,j'}, v'_{i,j'})$ . Then, again by Lemma 3.2 and the fact that each vertex in  $U'_i$  (in particular  $u'_{i,j'}$ ) has to belong to some path  $\pi(v) \in \mathcal{W}$ ,

$v \in U_i \setminus \{u\}$ , there exists in  $\mathcal{W}$  another inconsistent path  $\pi(u_{i,j''}) = (u_{i,j''}, u_i, u'_{i,j}, v_{i,j}, v'_{i,j})$  for some  $j'' \in I_i \setminus \{j\}$ . Now, we modify the solution by substituting  $\pi(u_{i,j}) := PV_{i,j}$  and  $\pi(u_{i,j''}) := (u_{i,j''}, u_i, u'_{i,j}, v_{i,j}, v'_{i,j})$ . Clearly, the two new paths still share vertex  $u_i$  and cover exactly the same vertices as the two original ones. Therefore, the modified set of paths is also a solution to the ST problem, moreover, with less number of inconsistent paths.

By repeating the above replacement argument a finite number of times, if ever needed, we obtain the desired solution satisfying Property (a).

(b) Consider any solution, say  $\mathcal{W}$ , to the ST problem for  $D_{SC}$  satisfying already proved Property (a). Consider any  $t \in [k]$ . If  $\pi(z_t) \in \mathcal{W}$  ends at a vertex of some horizontal path  $PH_l$  for some  $l \in [m]$  and  $\pi(z_t) \neq (z_t, z) \circ PH_l$ , then we just extend  $\pi(z_t)$  to have  $\pi(z_t) = (z_t, z) \circ PH_l$ . Otherwise, by the construction of  $D_{SC}$ , we must have  $\pi(z_t) = (z_t, z) \circ P \circ (v_{i,l}, v'_{i,l})$  for some  $i \in \mathcal{U}$ , where  $P$  is a subpath of  $PH_l$  for some  $l \in [m]$ . By the choice of  $\mathcal{W}$ , the arc  $(v_{i,l}, v'_{i,l})$  belongs to the (consistent) path  $\pi(u_{i,j}) = PV_{i,j}$  for some  $j \in I_i$ , and hence the path  $\pi(z_t)$  can be replaced by:  $\pi(z_t) := (z_t, z) \circ PH_l$ . Since  $P$  is a subpath of the new path  $\pi(z_t)$ , we conclude that the new set of paths is also a solution to the ST problem for  $D_{SC}$ , and moreover, it maintains the property that  $\pi(u_{i,j}) = PV_{i,j}$  for each  $i \in \mathcal{U}$  and  $j \in I_i$ .

Therefore, by repeating the above replacement argument a finite number of times, if ever needed, we obtain the desired solution satisfying both Properties (a) and (b)  $\square$

Now, we are going to prove our final lemma.

**Lemma 3.4.** *If there is a solution to the ST problem for  $D_{SC} = (\mathcal{V}, \mathcal{A}, \mathcal{F}, B)$ , then there exists a set cover of size  $k$  for the set system  $(\mathcal{U}, \mathcal{S})$ .*

**Proof.** By Lemma 3.3, any solution to the ST problem can be modified to be composed of the following paths:  $\pi(u_{i,j}) = PV_{i,j}$  for each  $i \in \mathcal{U}$  and  $j \in I_i$ , and  $\pi(z_t) = (z_t, z) \circ PH_{\xi(t)}$  for each  $t \in [k]$ , where  $\xi(t) \in [m]$ . Now, we claim that the set  $\mathcal{X}^l = \{\xi(1), \dots, \xi(k)\}$  is a set cover solution for the instance  $(\mathcal{U}, \mathcal{S}, k)$ . Indeed, since our solution to the ST problem is valid, for each  $i \in \mathcal{U}$  there exists  $t \in [k]$  such that  $v_{i,j}$  is a vertex of  $\pi(z_t)$ , since otherwise, in the underlying simple graph induced by our solution, no vertex in  $C_i$  is connected by a path to the vertex  $z$ . Thus,  $i \in S_{\xi(t)}$  which completes the proof.  $\square$

Note that the ST problem is clearly in NP and, as already observed, the construction of  $D_{SC}$  is polynomial in the input size to the SC problem. Hence, by combining Lemmas 3.1 and 3.4, we obtain the following result.

**Theorem 3.1.** *The ST problem is strongly NP-complete even for directed acyclic graphs  $D = (\mathcal{V}, \mathcal{A}, \mathcal{F}, B)$  with  $\mathcal{F} = \mathcal{V}$  and  $B(v) = 1$  if  $v$  is a source vertex in  $D$  and  $B(v) = 0$  otherwise.  $\square$*

*No pre-specified positions of snow ploughs* We claim that the Snow Team problem with Unspecified snow team bases is also NP-complete. The reduction is exactly the same as for the ST problem. All we need is to observe that if facilities are located at all vertices of the input digraph, then the number of snow ploughs sufficient to solve the STU problem is bounded from below by the number of source vertices in the digraph, since there must be at least one snow plough at each of its source vertices. Furthermore, without loss of generality we may assume that in any feasible solution of  $k$  walks, all snow ploughs are initially located at source vertices. Since in the digraph  $D_{SC}$  constructed for the proof of Theorem 3.1, we have  $B(v) = 1$  if  $v$  is a source vertex in  $D_{SC}$  and  $B(v) = 0$  otherwise, we may conclude with the following corollary.

**Corollary 3.1.** *The STU problem is strongly NP-complete even for directed acyclic graphs  $D = (\mathcal{V}, \mathcal{A}, \mathcal{F})$  with  $\mathcal{F} = \mathcal{V}$  and  $k$  being equal to the number of source vertices in  $D$ .  $\square$*

Since by setting  $\mathcal{F} = \mathcal{V}$ , the STU problem becomes just the Agent Clearing Tree problem (ACT) studied in [22], we immediately obtain the following corollary resolving the open problem of the complexity status of ACT posed in [22].

**Corollary 3.2.** *The ACT problem is NP-complete.  $\square$*

#### 4. Open problem

In all of our variants of the Snow Team problem, we assumed that a snow plough can traverse arbitrary number of arcs. However, from a practical point of view, it is more natural to assume that each snow plough, called an *s-plough*, can traverse and clear only the fixed number  $s$  of arcs [25]. Observe that in this case, the key Lemma 2.2 does not hold, which immediately makes our algebraic approach unfeasible for the Snow Team problem with *s-ploughs*, so this variant requires further studies.

## Acknowledgment

This research has been partially supported by National Science Centre (Poland) grant number 2015/17/B/ST6/01887, and Swedish Research Council, grant number 2017-03750. The authors thank the anonymous reviewers for their careful reading of our manuscript and their many insightful comments and suggestions.

## References

- [1] A. Abdi, A.E. Feldmann, B. Guenin, J. Könnemann, L. Sanità, Lehman's theorem and the directed Steiner tree problem, *SIAM J. Discrete Math.* 30 (1) (2016) 141–153.
- [2] N. Beerenwinkel, S. Beretta, P. Bonizzoni, R. Dondi, Y. Pirola, Covering pairs in directed acyclic graphs, *Comput. J.* 58 (7) (2015) 1673–1686.
- [3] A. Björklund, T. Husfeldt, P. Kaski, M. Koivisto, Narrow sieves for parameterized paths and packings, *J. Comput. Syst. Sci.* 87 (2017) 119–139.
- [4] A. Björklund, T. Husfeldt, N. Taslamani, Shortest cycle through specified elements, in: *Twenty-Third Annual ACM–SIAM Symposium on Discrete Algorithms (SODA 2012)*, 2012, pp. 1747–1753.
- [5] A. Björklund, P. Kaski, L. Kowalik, Constrained multilinear detection and generalized graph motifs, *Algorithmica* 74 (2) (2016) 947–967.
- [6] P. Borowiecki, D. Dereniowski, P. Pralat, Brushing with additional cleaning restrictions, *Theor. Comput. Sci.* 557 (2014) 76–86.
- [7] D. Bryant, N. Francetic, P. Gordinowicz, D.A. Pike, P. Pralat, Brushing without capacity restrictions, *Discrete Appl. Math.* 170 (2014) 33–45.
- [8] M. Charikar, C. Chekuri, T.-Y. Cheung, Z. Dai, A. Goel, S. Guha, M. Li, Approximation algorithms for directed Steiner problems, *J. Algorithms* 33 (1) (1999) 73–91.
- [9] R.H. Chitnis, H. Esfandiari, M. Hajiaghayi, R. Khandekar, G. Kortsarz, S. Seddighin, A tight algorithm for strongly connected Steiner subgraph on two terminals with demands (extended abstract), in: *9th International Symposium on Parameterized and Exact Computation (IPEC 2014)*, 2014, pp. 159–171.
- [10] A.E. Feldmann, D. Marx, The complexity landscape of fixed-parameter directed Steiner network problems, in: *43rd International Colloquium on Automata, Languages, and Programming (ICALP 2016)*, 2016, pp. 27:1–27:14.
- [11] F.V. Fomin, D. Lokshtanov, V. Raman, S. Saurabh, B.V.R. Rao, Faster algorithms for finding and counting subgraphs, *J. Comput. Syst. Sci.* 78 (3) (2012) 698–706.
- [12] Z. Friggstad, J. Könnemann, Y. Kun-Ko, A. Louis, M. Shadravan, M. Tulsiani, Linear programming hierarchies suffice for directed Steiner tree, in: *17th International Conference on Integer Programming and Combinatorial Optimization (IPCO 2014)*, 2014, pp. 285–296.
- [13] H.N. Gabow, S.N. Maheshwari, L.J. Osterweil, On two problems in the generation of program test paths, *IEEE Trans. Softw. Eng.* 2 (3) (1976) 227–231.
- [14] M.R. Garey, D.S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*, W.H. Freeman & Co., New York, NY, USA, 1979.
- [15] S. Gaspers, M.-E. Messinger, R.J. Nowakowski, P. Pralat, Parallel cleaning of a network with brushes, *Discrete Appl. Math.* 158 (5) (2010) 467–478.
- [16] M. Jones, D. Lokshtanov, M.S. Ramanujan, S. Saurabh, O. Suchý, Parameterized complexity of directed Steiner tree on sparse graphs, in: *21st European Symposium on Algorithms (ESA 2013)*, 2013, pp. 671–682.
- [17] P. Kolman, O. Pangrác, On the complexity of paths avoiding forbidden pairs, *Discrete Appl. Math.* 157 (13) (2009) 2871–2876.
- [18] I. Koutis, Faster algebraic algorithms for path and packing problems, in: *35th International Colloquium on Automata, Languages and Programming (ICALP 2008)*, 2008, pp. 575–586.
- [19] I. Koutis, R. Williams, Limits and applications of group algebras for parameterized problems, *ACM Trans. Algorithms* 12 (3) (2016) 31.
- [20] J. Kováč, Complexity of the path avoiding forbidden pairs problem revisited, *Discrete Appl. Math.* 161 (10–11) (2013) 1506–1512.
- [21] B. Laekhanukit, Approximating directed Steiner problems via tree embedding, in: *43rd International Colloquium on Automata, Languages, and Programming (ICALP 2016)*, 2016, pp. 74:1–74:13.
- [22] C. Levcopoulos, A. Lingas, B.J. Nilsson, P. Żyliński, Clearing connections by few agents, in: *7th International Conference Fun with Algorithms (FUN 2014)*, 2014, pp. 289–300.
- [23] M.-E. Messinger, R.J. Nowakowski, P. Pralat, Cleaning a network with brushes, *Theor. Comput. Sci.* 399 (3) (2008) 191–205.
- [24] M.-E. Messinger, R.J. Nowakowski, P. Pralat, Cleaning with brooms, *Graphs Comb.* 27 (2) (2011) 251–267.
- [25] S.C. Ntafos, T. Gonzalez, On the computational complexity of path cover problems, *J. Comput. Syst. Sci.* 29 (2) (1984) 225–242.
- [26] S.C. Ntafos, S.L. Hakimi, On path cover problems in digraphs and applications to program testing, *IEEE Trans. Softw. Eng.* 5 (5) (1979) 520–529.
- [27] S.C. Ntafos, S.L. Hakimi, On structured digraphs and program testing, *IEEE Trans. Comput. C-30* (1) (1981) 67–77.
- [28] R. Otter, The number of trees, *Ann. Math. Second Ser.* 49 (3) (1948) 583–599.
- [29] Ondrej Suchý, On directed Steiner trees with multiple roots, in: *42nd International Workshop on Graph-Theoretic Concepts in Computer Science (WG 2016)*, 2016, pp. 257–268.
- [30] D. Watel, M.-A. Weisser, C. Bentz, D. Barth, Directed Steiner tree with branching constraint, in: *20th International Computing and Combinatorics Conference (COCOON 2014)*, 2014, pp. 263–275.
- [31] D. Watel, M.-A. Weisser, C. Bentz, D. Barth, Directed Steiner trees with diffusion costs, *J. Comb. Optim.* 32 (4) (2016) 1089–1106.
- [32] R. Williams, Finding paths of length  $k$  in  $O^*(2^k)$  time, *Inf. Process. Lett.* 109 (6) (2009) 315–318.
- [33] M. Zientara, Personal communication, 2016.
- [34] L. Zosin, S. Khuller, On directed Steiner trees, in: *Thirteenth Annual ACM–SIAM Symposium on Discrete Algorithms (SODA 2002)*, 2002, pp. 59–63.