Available at
www.ElsevierMathematics.com
POWERED BY SCIENCE DIRECT
Discrete Applied Mathematics 138 (2004) 289–301
DISCRETE APPLIED MATHEMATICS
www.elsevier.com/locate/dam

Interval data minmax regret network optimization problems

Igor Averbakh[a], Vasilij Lebedev[b]

[a] Division of Management, University of Toronto at Scarborough, 1265 Military Trail, Scarborough, Ont. Canada M1C 1A4

[b] Department of Mathematics, Volgograd State University, Volgograd 400062, Russia

Received 15 October 2001; received in revised form 24 May 2002; accepted 5 June 2003

Abstract

We consider the minimum spanning tree and the shortest path problems on a network with uncertain lengths of edges. In particular, for any edge of the network, only an interval estimate of the length of the edge is known, and it is assumed that the length of each edge can take on any value from the corresponding interval of uncertainty, regardless of the values taken by the lengths of other edges. It is required to find a minmax regret solution. We prove that both problems are NP-hard even if the bounds of all intervals of uncertainty belong to $\{0,1\}$. The interval data minmax regret shortest path problem is NP-hard even if the network is directed, acyclic, and has a layered structure. We show that the problems are polynomially solvable in the practically important case where the number of edges with uncertain lengths is fixed or is bounded by the logarithm of a polynomial function of the total number of edges. We discuss implications of these results for the general theory of interval data minmax regret combinatorial optimization.

© 2003 Elsevier B.V. All rights reserved.

Keywords: Minmax regret optimization; Spanning tree; Shortest path

1. Introduction

Minmax regret optimization deals with optimization problems where the objective function is uncertain at the time of solving the problem, and it is required to find a feasible solution that is ε-optimal for any possible scenario (realization of the objective function), with ε as small as possible. The book [5] gives the state-of-art in

E-mail addresses: averbakh@scar.utoronto.ca, averbakh@utsc.utoronto.ca (I. Averbakh), lebedevvn@mail.ru (V. Lebedev).

0166-218X/$ - see front matter © 2003 Elsevier B.V. All rights reserved.
doi:10.1016/S0166-218X(03)00462-1

minmax regret combinatorial optimization (MRCO) up to 1997 and provides a comprehensive discussion of the motivation for the minmax regret approach and various aspects of applying it in practice. Minmax regret solutions are sometimes called *robust* solutions [5], although there are different concepts of robustness in the literature (e.g., [6,8]).

A specific objective function is typically defined by a vector of numerical parameters; in network optimization problems (e.g. the shortest path or minimum spanning tree problems), such parameters are the lengths of edges of the network. Thus, we assume that each scenario s can be considered as a vector in R^m, where m is the number of relevant numerical parameters. Two natural ways of describing the set of all possible scenarios S have been considered in the literature. In the *discrete-scenario case*, S is assumed to be finite and described by explicitly listing all vectors $s \in S$. In the *interval data case*, it is assumed that for each numerical parameter, only lower and upper bounds for the value of this parameter are known, and the parameter can take on any value between these bounds, regardless of values taken by other parameters. Thus, in this case S is the Cartesian product of the intervals of uncertainty for the parameters.

Complexity of discrete-scenario MRCO problems has been studied quite extensively [5]. A general observation is that minmax regret versions of most classical combinatorial optimization problems are NP-hard in the case of discrete-scenario representation of uncertainty [5]. On the contrary, little is known about complexity of interval data MRCO problems. A natural conjecture would be that if a MRCO problem is NP-hard in the discrete-scenario case, it is also NP-hard in the interval data case. However, this is not true: Averbakh [3] showed that there are MRCO problems that are NP-hard in the discrete-scenario case but are polynomially solvable in the interval data case. (E.g., the problem of selecting p objects of the smallest total weight out of a total of $n > p$ objects, with uncertainty in the weights of the objects [3].)

In this paper, we study the interval data minmax regret minimum spanning tree and shortest path problems. Such problems have already been considered in the literature. Kouvelis and Yu [5] studied the discrete-scenario versions of the problems and proved that they are NP-hard even if there are only two scenarios and strongly NP-hard if the number of scenarios is unbounded. Yaman et al. [9,10] studied some structural properties of the interval data minmax regret minimum spanning tree problem and the longest path problem in a directed acyclic network. Montemanni and Gambardella [7] developed a branch-and-bound algorithm for the interval data minmax regret spanning tree problem. The complexity of the problems in the interval data case has remained open so far.

Our main results are:

(1) The interval data minmax regret minimum spanning tree problem is NP-hard even if all intervals of uncertainty are equal to $[0, 1]$.
(2) The interval data minmax regret shortest path problem on an undirected network is NP-hard even if all bounds of all intervals of uncertainty belong to $\{0, 1\}$.
(3) The interval data minmax regret shortest path problem on a directed network is NP-hard even if the network is acyclic, has layered structure, and all bounds of intervals of uncertainty belong to $\{0, 1\}$.

(4) If the number of edges with uncertain lengths is fixed or is bounded by the logarithm of a polynomial function of the total number of edges, the above problems (and any interval data minmax regret minisum network optimization problems that are polynomially solvable without uncertainty) are polynomially solvable.

The choice of the problems to study is not random. The minimum spanning tree and the shortest path problems are the basic and the simplest minisum network optimization problems. Thus, NP-hardness of their interval data minmax regret versions can be considered as an evidence that sufficiently general polynomially solvable classes of interval data MRCO problems are unlikely to be found among minisum network optimization problems if no bounds are imposed on the number of edges with uncertain lengths. Since our NP-hardness results hold for very simply structured intervals of uncertainty, polynomial solvability cannot be achieved by any reasonable additional assumptions about the structure of intervals of uncertainty. Also, these results provide an insight about applicability of some classical combinatorial optimization approaches to interval data MRCO problems. First, NP-hardness of the interval data minimum spanning tree problem shows that in interval data minmax regret optimization, matroidal structure of the set of feasible solutions is not sufficient for polynomial solvability. Second, NP-hardness of the interval data minmax regret shortest path problem shows that combinatorial optimization problems that can be solved in polynomial time by dynamic programming without uncertainty may not be polynomially solvable in the interval data minmax regret version.

The positive result (4) demonstrates that interval data minmax regret network optimization models can be efficiently used in situations where the number of edges with uncertain lengths is sufficiently small. Also, this result shows that interval data minmax regret network optimization problems are in a sense more tractable than their discrete-scenario counterparts (the discrete-scenario minmax regret minimum spanning tree and shortest path problems are NP-hard even if there are only two scenarios [5]). We note also that the negative results (1)–(3) pertain to the case of problems with *minisum* objective function. As shown in [2], network optimization problems with a *minimax* (*bottleneck*) type of objective function are polynomially solvable in the interval data minmax regret version whenever their classical versions (without uncertainty) are polynomially solvable.

2. Preliminaries

A generic combinatorial optimization problem can be formulated as follows:

Problem OPT. Minimize $\{F(X) | X \in A\}$, where A is a set of feasible solutions (a finite or compact set) and $F(\cdot)$ is a function defined on A (that has the necessary continuity properties if A is an infinite compact set so that an optimum always exists). Suppose that there is uncertainty in the objective function, that is, it is known only that $F(\cdot)$ is a member of a family of functions $\{F_s(\cdot), s \in S\}$ for some *set of scenarios* S. The set A

is scenario-independent. Let F_s^* denote the optimum objective value for the following problem:

Problem OPT(s)**.** Minimize $\{F_s(X)|X \in A\}$.
For any $X \in A$ and $s \in S$, the value $R(s,X) = F_s(X) - F_s^*$ is called the *regret for X under scenario s*. For any $X \in A$, the value

$$Z(X) = \max_{s \in S} R(s,X) \tag{1}$$

is called a *worst-case regret* for X. The minmax regret version of Problem OPT is

Problem ROB. Minimize $\{Z(X)|X \in A\}$.

The acronym "ROB" in Problem ROB refers to robustness.
For any $X, Y \in A$, let

$$r(X,Y) = \max_{s \in S} (F_s(X) - F_s(Y)). \tag{2}$$

Then $Z(X)$ can be written alternatively as

$$Z(X) = \max_{Y \in A} r(X,Y) \tag{3}$$

or

$$Z(X) = \max_{s \in S} \max_{Y \in A} (F_s(X) - F_s(Y)). \tag{4}$$

An optimal solution to the right-hand side of (3) is called a *worst-case alternative* for X. An optimal solution to the right-hand side of (1) is called a *worst-case scenario* for X. An optimal solution to the right-hand side of (2) is called a *worst-case scenario for X with respect to Y*. An optimal solution $(\hat{s}, \hat{Y})$ to the right-hand side of (4) is called a *worst-case pair for X*. Observe that if $(\hat{s}, \hat{Y})$ is a worst-case pair for X, then $\hat{s}$ is a worst-case scenario and $\hat{Y}$ is a worst-case alternative for X.

Let $G = (V,E)$ be an undirected connected graph with V being the set of nodes and E being the set of edges, $|V| = n$, $|E| = m$. Suppose that for every edge $e \in E$, two nonnegative integer numbers c_e^+, c_e^- are given, $c_e^- \leqslant c_e^+$. The numbers c_e^-, c_e^+ represent the lower and the upper bounds on the length of edge e. It is assumed that the length of edge e can take on any real value from its interval of uncertainty $[c_e^-, c_e^+]$, regardless of the values taken by the lengths of other edges, so the set of scenarios S is the Cartesian product of the intervals of uncertainty $[c_e^-, c_e^+], e \in E$. For any integers k, t, $k \leqslant t$, let $[k:t]$ denote the set of integers between k and t (including k, t).

3. The interval data minmax regret minimum spanning tree problem

In this section, we assume that the set of feasible solutions A is the set of all spanning trees of the network G. For any scenario $s = \{c_e^{(s)}, e \in E\} \in S$ and any $X \in A$, we assume $F_s(X) = \sum_{e \in X} c_e^{(s)}$. Problems OPT($s$) and ROB in this case will be referred to as **Problem OPT.TREE**(s) and **Problem ROB.TREE**, respectively. It is well known that Problem OPT.TREE(s) can be solved in strongly polynomial time (for a survey, see [1]). The main result of the section is

Theorem 1. *Problem ROB.TREE is NP-hard even if all intervals of uncertainty are equal to* $[0, 1]$.

Corollary. *Problem ROB.TREE is strongly NP-hard.*

In the proof of Theorem 1 we use the following auxiliary problems.

Problem P1. Given is a connected undirected graph G'. Find the maximum number of connected components that can be obtained from G' by removing the edges of some spanning tree of G'.

Problem P2 (Exact cover by 3-sets). Given is a finite set B, $|B| = 3q$ for some integer q, and a collection T of 3-element subsets of B (triples).

Question. Does T contain an exact cover for B, i.e. a subcollection $T' \subseteq T$ such that every element of B occurs in exactly one member of T'?

Problem P2 is known to be NP-complete (see [4]). In order to prove Theorem 1, we first show that Problem P1 is NP-hard using a reduction from Problem P2; then we show that Problem ROB.TREE is NP-hard even if all intervals of uncertainty are equal to $[0, 1]$, using a reduction from Problem P1.

Lemma 1. *Problem P1 is NP-hard.*

Proof. Consider an instance $\langle B, T \rangle$ of Problem P2. Assume that $q \geqslant 2$. The corresponding instance of Problem P1 (graph G') is obtained as follows. For any $b \in B$, we create two nodes v_b, u_b; for any triple $t \in T$, we create a node w_t. Then, $\{v_b, u_b, w_t | b \in B, t \in T\}$ is the set of nodes of the graph G'. The set of $2|B|$ nodes that correspond to elements of B is denoted by $V(B)$; the set of $|T|$ nodes that correspond to elements of T is denoted by $V(T)$.

The set of edges of graph G' is obtained as follows. First, we include all possible edges joining nodes from $V(B)$ (that is, the subgraph of G' induced by the set of nodes $\{v_b, u_b | b \in B\}$ is complete). Second, for any triple $t = (b_1, b_2, b_3) \in T$, we include edges (w_t, v_{b_1}), (w_t, v_{b_2}), (w_t, v_{b_3}), (w_t, u_{b_1}), (w_t, u_{b_2}), (w_t, u_{b_3}).

To illustrate constructing graph G' from sets B and T, consider a small example. Suppose that $B = \{b_1, b_2, b_3, b_4\}$, $T = \{t_1, t_2\}$, where $t_1 = \{b_1, b_2, b_3\}$, $t_2 = \{b_2, b_3, b_4\}$. (Here, $|B|$ is not a multiple of 3, but this is not important for illustration purposes). Then, the corresponding graph G' is depicted in Fig. 1. The rectangle denotes a complete subgraph.

Lemma 1 follows from the following statement.

Statement 1. The original instance of Problem P2 has answer "Yes" if and only if the optimal objective value for the obtained instance of Problem P1 is $q + 1$.

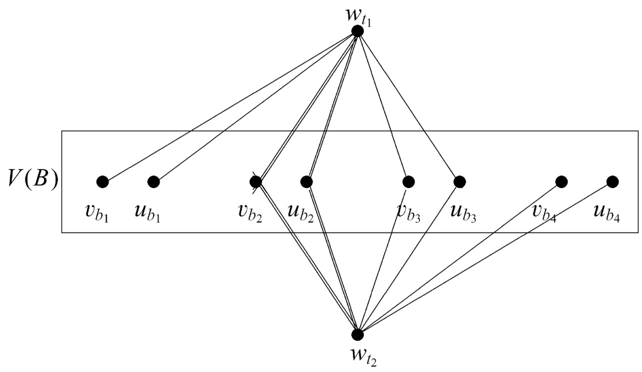

Fig. 1. Illustration for the proof of Lemma 1.

Let us prove Statement 1. Let μ denote the optimal objective value for the obtained instance of Problem P1. Let G'' denote the graph obtained from G' by removing the edges of an optimal spanning tree; G'' has μ connected components.

Observation 1. *If for some triples $t_1, t_2 \in T$, $\{w_{t_1}\}$, $\{w_{t_2}\}$ are connected components of G'', then the triples t_1 and t_2 do not have common elements.*

Indeed, otherwise the set of removed edges would contain a cycle, which is impossible; such a cycle is shown in Fig. 1 by the doubled edges.

Observation 2. *If $\mu \geqslant 3$, then all $6q$ nodes from $V(B)$ belong to the same connected component of G'' which will be called the* main component.

Indeed, the nodes from $V(B)$ induce a complete subgraph of G'. If the number of connected components of G'' that contain nodes from $V(B)$ is greater than 2, then the set of removed edges must contain a cycle, which is impossible. Suppose now that the number of connected components of G'' that contain nodes from $V(B)$ is equal to 2. If each of the two components contains at least two nodes from $V(B)$, then it is easy to see that the set of removed edges must contain a cycle, which is impossible. Thus, we can assume that one of these two components contains only one node from $V(B)$ (say, node v). Let R denote the other component that contains nodes from $V(B)$. Then we notice that all nodes $w_t, t \in T$, must belong to R. Indeed, if for some triple $t' \in T$ vertex $w_{t'}$ does not belong to R, then the set of removed edges would contain a cycle $(v, v_b, w_{t'}, u_b, v)$ for some $b \in t'$, which is impossible. Then graph G'' has exactly two connected components, which is impossible since we assumed $\mu \geqslant 3$.

Observation 3. *If $\mu \geqslant q+1$, then $\mu = q+1$, and graph G'' consists of the main component and q single-node components $\{w_t\}$, $t \in T'$, where T' is an exact cover for B.*

Indeed, from the assumption $q \geqslant 2$ and Observation 2 it follows that all nodes of $V(B)$ belong to the same component (main component). Thus, each of the remaining $\mu - 1$ components consists of a single node w_t for some $t \in T$. Taking into account Observation 1, we have that these $\mu - 1$ single-node components correspond to triples that do not have common elements; thus, $\mu - 1 \leqslant q$, and since $\mu \geqslant q + 1$ we have $\mu = q + 1$. Therefore, G'' has q single-node components $\{w_t\}$, $t \in T'$, where T' is an exact cover for B.

Suppose $\mu = q + 1$. Then Observation 3 implies that the answer to the original instance of Problem P2 is "Yes".

Suppose now that there exists an exact cover $T' \subset T$ for B. Let E' be the set of edges that connect nodes $w_t, t \in T'$, with nodes from $V(B)$. E' does not contain a cycle, because no two triples from T' have common elements. Therefore, by adding $q - 1$ edges of G' to E' we can obtain a spanning tree. Deleting this spanning tree from G' results in a graph with at least $q + 1$ components; then, $\mu = q + 1$ according to Observation 3. This completes the proof of Statement 1.

Lemma 1 is proven. □

Now let us prove Theorem 1 by means of a reduction from Problem P1. Suppose we are given an instance $G' = (V, E)$ of Problem P1, $|V| = n$, $|E| = m$. The corresponding instance of Problem ROB.TREE is obtained by using the same graph $G = G' = (V, E)$ and setting the intervals of uncertainty for the lengths of all edges equal to $[0, 1]$. The following statement directly implies Theorem 1.

Statement 2. If z^* is the optimal objective value for the obtained instance of Problem ROB.TREE, then $n - z^*$ is the optimal objective value for the original instance of Problem P1.

Statement 2 immediately follows from the following.

Statement 3. For any spanning tree X of graph G, if the graph G'' obtained by removing the edges of X from G has k connected components, then $Z(X) = n - k$, where $Z(X)$ is defined in (1).

Let us prove Statement 3. Let s_X denote the scenario obtained by setting the lengths of the edges from X equal to 1, and setting the lengths of other edges equal to 0; clearly s_X is a worst-case scenario for X.

Observe that since removing the edges of X decomposes graph G into k connected components, then any spanning tree $Y \in A$ has at least $k - 1$ edges in common with X. Taking into account that any spanning tree of G has $n - 1$ edges, we have that for any $Y \in A$, $F(s_X, X) - F(s_X, Y) \leqslant (n - 1) - (k - 1) = n - k$ and therefore $Z(X) \leqslant n - k$.

Now, let us obtain a spanning tree $Y' \in A$ as follows. Take arbitrary spanning trees in the connected components of graph G'', and connect these components using $k - 1$ edges of X so that no cycles are formed (this is possible since graph G is connected). The spanning tree $Y' \in A$ obtained in this way has exactly $k - 1$ edges of X; therefore, $F(s_X, X) - F(s_X, Y') = (n - 1) - (k - 1) = n - k$, and thus $Z(X) \geqslant n - k$ (see (4)).

Combining the obtained inequalities $Z(X) \leqslant n - k$ and $Z(X) \geqslant n - k$, we have $Z(X) = n - k$. Statement 3 is proved. This concludes the proof of Theorem 1.

Remark. For a given family F of sets S_i, a *transversal* is a set $\hat{S}$ that has nonempty intersection with all S_i. A *maximum* transversal is a transversal that maximizes the minimum intersection with a member of F. It follows from the proof that Problem P1, and, therefore, Problem ROB.TREE with $[0,1]$ uncertainty intervals, is equivalent to finding a maximum transversal of the family of all spanning trees which is also a spanning tree. A somewhat similar combinatorial idea will also be used in the next section.

4. The interval data minmax regret shortest path problem

Suppose that $a, b \in V$ are two fixed nodes of the network G. In this section, we assume that the set A of feasible solutions is the set of all simple paths in G from a to b. For any scenario $s = \{c_e^{(s)}, e \in E\}$ and any $X \in A$, we assume $F_s(X) = \sum_{e \in X} c_e^{(s)}$. Problems OPT($s$) and ROB in this case are referred to as **Problem OPT.PATH(s)** and **Problem ROB.PATH**, respectively. It is well known that Problem OPT.PATH(s) can be solved in strongly polynomial time (for a survey, see [1]). The main result of the section is

Theorem 2. *Problem ROB.PATH is NP-hard even if all bounds of the intervals of uncertainty belong to* $\{0,1\}$.

Corollary. *Problem ROB.PATH is strongly NP-hard.*

Let us prove Theorem 2. We use the following auxiliary problem:

Problem P3 (Hamiltonian path). Given is an undirected connected graph H.

Question. Does H contain a Hamiltonian path, that is, a path that visits each node exactly once?

Problem P3 is known to be NP-complete (see [4]).

We prove Theorem 2 by means of a reduction from Problem P3.

Let $H = (N, L)$ be an instance of Problem P3, where N is the set of nodes, $|N| = p$, and L is the set of edges of graph H, $|L| = q$. We assume $p \geqslant 2$. The corresponding instance of Problem ROB.PATH (defined by graph $G = (V, E)$, selected nodes $a, b \in V$—the endpoints of feasible paths, and intervals of uncertainty $[c_e^-, c_e^+]$ for all edges $e \in E$) is obtained as follows. The set of nodes V consists of nodes a, b, and $2p$ copies of the set N that will be called *blocks* (thus, $|V| = 2 + 2p^2$). Let N_i denote block i. Thus, $V = (\bigcup_{i \in [1:2p]} N_i) \cup \{a, b\}$. For any $v \in N$, v_i denotes the copy of v from block i.

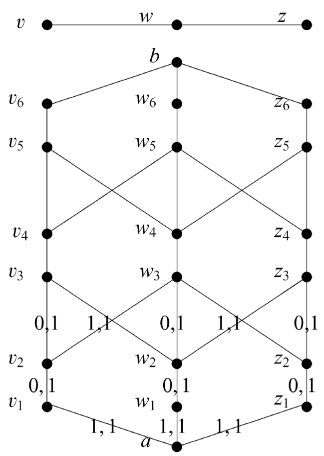

Fig. 2. Illustration for the proof of Theorem 2.

The set of edges E of graph G consists of:

(1) Edges (a,u), $u \in N_1$.
(2) Edges (b,u), $u \in N_{2p}$.
(3) Edges (v_i, v_{i+1}) for all $v \in N$, $i \in [1:2p-1]$. These edges are referred to as *vertical edges*, and the set of all vertical edges is denoted by $\tilde{E}$.
(4) Edges (v_{2i}, w_{2i+1}), (w_{2i}, v_{2i+1}) for all edges $(v,w) \in L$ of graph H, and all $i \in [1:p-1]$.

Edges of E that are not vertical edges are called *diagonal edges*. Intervals of uncertainty for vertical edges are equal to $[0,1]$, and for diagonal edges are equal to $[1,1]$. Let $\tilde{D}$ denote the set of diagonal edges not incident to a or b. For any $v \in N$, the path $(a, v_1, v_2, \ldots, v_{2p}, b)$ is called the *v-tunnel*.

Observe that nodes of "even" blocks $N_{2i}, i \in [1:p]$ can be reached from nodes of the "previous" blocks $N_{2i-1}, i \in [1:p]$ only via vertical edges.

An example of a graph H (with three nodes) and the corresponding graph G is depicted in Fig. 2.

Theorem 2 follows from the following statement.

Statement 4. The original instance of Problem P3 has answer "Yes" if and only if the optimum objective value for the obtained instance of Problem ROB.PATH is $2p-2$.

Proof. For a (simple) $a-b$ path X, let $|X|$ stand for the number of its edges, $\hat{X}$ for its maximal subpath from N_1 to N_{2p}, and s_X for the scenario taking the upper bounds of the uncertainty intervals on the edges of X and the lower bounds on the rest. Clearly s_X is a worst-case scenario for X, and $F_{s_X}(X) = |X| = |\hat{X}| + 2 \geqslant 2p+1$. Also, for another $a-b$ path Y, define $F(X,Y)=F_{s_X}(X)-F_{s_X}(Y)$; then

$$F(X,Y)=|\hat{X}|-|Y\cap\tilde{D}|-|Y\cap X\cap\tilde{E}|. \tag{5}$$

Suppose H has a Hamiltonian path, with the sequence of nodes $(x^1,x^2,\dots,x^p)$. Then $(a,x_1^1,x_2^1,x_3^2,x_4^2,\dots,x_{2p-1}^p,x_{2p}^p,b)$ forms path X, and $|\hat{X}|=2p-1$. Each tunnel T has no edge in $\tilde{D}$ and exactly one vertical edge in common with X. Hence $F(X,T)=2p-2$, by (5). Any nontunnel $a-b$ path Y contains at least one edge in $\tilde{D}$, implying $F(X,Y)\leqslant 2p-2$. Thus, $Z(X)=2p-2$, and the optimal objective value for the problem is at most $2p-2$. (In fact exactly $2p-2$, as follows from the rest of the proof.)

Now consider an optimal $a-b$ path X for a generic instance that has the optimal objective value not greater than $2p-2$. Let T be a tunnel with $k=|\hat{T}\cap\hat{X}|$ $(=|T\cap X\cap\tilde{E}|)$ minimum. Assume $F(X,T)\leqslant 2p-2$ (otherwise $Z(X)\geqslant 2p-1$ and the optimal objective value is at least $2p-1$). Since $F(X,T)=|\hat{X}|-k$ and $|\hat{X}|\geqslant 2p-1$, we have $k\geqslant 1$. Then each of the p tunnels intersects $\hat{X}$ by at least $k\geqslant 1$ (vertical) edges, these edges in $\hat{X}$ are different, and there are at least $p-1$ diagonal edges in $\hat{X}$ that connect its intersections with the tunnels. We have

$$2p-2\geqslant|\hat{X}|-k\geqslant(kp+p-1)-k=(k+1)(p-1)\geqslant 2p-2.$$

So equality holds throughout, implying that $k=1$, $|\hat{X}|=2p-1$, $F(X,T)=2p-2$, $Z(X)=2p-2$, and that X has exactly one vertical edge of the form $\{v_j,v_{j+1}\}$ for each $v\in N$ and exactly one diagonal edge of the form $\{u_{2i},v_{2i+1}\}$ for each $i=1,\dots,p-1$. Then the images in H of the diagonal edges of $\hat{X}$ form a Hamiltonian path, and the result follows. □

Theorem 2 is proven.

Remark. The proof of Theorem 2 can be used (almost without modifications) to prove NP-hardness of the interval data minmax regret shortest path problem on a *directed* network, even if the network is acyclic and has a special (layered) structure. Suppose that $G=(V,E)$ is a directed network with the set of nodes V and the set of arcs E and with selected nodes $a,b\in V$. Suppose that the set A of feasible solutions is the set of all simple directed paths in G from a to b. All other definitions and notation are exactly the same as in the undirected case. A directed network is called *acyclic* if it has no *directed* cycles. A directed network is called *layered* if its set of vertices V can be partitioned into nonoverlapping subsets $V_1,V_2,\dots,V_k$ (called blocks) such that each arc goes from a node of some block V_i to a node of the next block V_{i+1}. If a directed network is layered it is also acyclic. The regular shortest path problem on an acyclic directed network can be solved in linear time [1] by dynamic programming. We note that if in the proof of Theorem 2, for the network G of the constructed instance of Problem ROB.PATH we consider all edges as arcs directed from blocks with smaller indices to blocks with larger indices (assuming $N_0=\{a\}$, $N_{2p+1}=\{b\}$), the network will be layered, and we obtain a proof that Problem ROB.PATH is NP-hard in the case of directed acyclic layered networks. (In fact, in the directed case the reasoning is even simpler than in the undirected case because for the constructed instance of Problem ROB.PATH, any directed path from a to b has exactly $2p+1$ arcs).

In many applications (e.g. project management) it is important to find the *longest* path in a directed acyclic network. The regular longest path problem in a directed acyclic network can be solved in linear time [1] by dynamic programming. The interval data minmax regret longest path problem corresponds to Problem ROB.PATH where for any $X \in A$ and any $s = \{c_e^{(s)}, e \in E\} \in S$, $F_s(X) = -\sum_{e \in X} c_e^{(s)}$. NP-hardness of this problem in a directed acyclic layered network can be proved using the same reduction, with only one modification: Set the intervals of uncertainty for the diagonal arcs equal to $[0, 0]$ (instead of $[1, 1]$).

Let us formulate the results for the directed case as a theorem.

Theorem 3. *The interval data minmax regret shortest* (*or longest*) *path problem on a directed network is NP-hard even if the network has no directed cycles*, *has layered structure*, *and all bounds of intervals of uncertainty belong to* $\{0, 1\}$.

Remark. The considered problems remain NP-hard even in the special case where the lengths of uncertainty intervals are small with respect to "average" lengths of the corresponding edges, that is, when $2(c_e^+ - c_e^-)/(c_e^+ + c_e^-) \ll 1$ for all $e \in E$. For Problem ROB.TREE, this follows from the observation that all spanning trees have the same number of edges and therefore adding the same constant to all lower and upper bounds of uncertainty intervals does not change the optimal objective value. For Problem ROB.PATH the argument is similar, when we notice that in our proof of NP-hardness all relevant paths have the same number of edges.

5. Some polynomially solvable cases

In this section, we consider a generic minisum network optimization problem, that is, we assume that the set A of feasible solutions is a set of some subsets of E, and for any $X \in A$ and $s = \{c_e^{(s)}, e \in E\} \in S$, $F_s(X) = \sum_{e \in X} c_e^{(s)}$. (Clearly, the problems discussed in Sections 3 and 4 belong to this class.) The interval of uncertainty $[c_e^-, c_e^+]$ for an edge $e \in E$ is called *nondegenerate* if $c_e^- < c_e^+$, and is called *degenerate* if $c_e^- = c_e^+$. Let d be the number of nondegenerate intervals of uncertainty. In this section we derive some properties of the problems under consideration that allow us to solve Problem ROB in polynomial time whenever Problem OPT(s) is solvable in polynomial time and d is sufficiently small (asymptotically) with respect to the size m of the network; namely, when d is bounded by the logarithm of a polynomial function of m.

For any $X \in A$, let $s(X) = \{c_e^{s(X)}, e \in E\}$ denote the scenario defined as follows:

$$c_e^{s(X)} = \begin{cases} c_e^+ & \text{if } e \notin X, \\ c_e^- & \text{if } e \in X. \end{cases}$$

Lemma 2. *Consider an arbitrary* $X \in A$. *Let* $\tilde{X}$ *be an optimal solution to Problem* $OPT(s(X))$. *Then* $Z(\tilde{X}) \leqslant Z(X)$.

Proof. It is straightforward to observe that for any scenario $s \in S$, $F_s(X) - F_s(\tilde{X}) \geqslant F_{s(X)}(X) - F_{s(X)}(\tilde{X})$. Since $F_{s(X)}(X) \geqslant F_{s(X)}(\tilde{X})$ we have that for any $s \in S$, $F_s(X) \geqslant F_s(\tilde{X})$, which implies the statement of the lemma. $\square$

Corollary. *If $X^* \in A$ is an optimal solution to Problem ROB, then any optimal solution to Problem OPT$(s(X^*))$ is also an optimal solution to Problem ROB.*

A scenario $s \in S$ is called an *extreme* scenario if under this scenario the length of any edge $e \in E$ is equal to one of the endpoints of the corresponding interval of uncertainty. Let S' denote the set of extreme scenarios; then $|S'| = 2^d$. According to the corollary from Lemma 2, there exists an extreme scenario $s \in S'$ such that any optimal solution to Problem OPT(s) is also an optimal solution to Problem ROB. This justifies the following simple algorithm for solving Problem ROB. For each extreme scenario $s \in S'$, find an optimal solution $\tilde{X}(s)$ to Problem OPT(s), and compute value $Z(\tilde{X}(s))$. Then, the best of the obtained candidate solutions will be an optimal solution to Problem ROB. Since computing value $Z(X)$ for any $X \in A$ amounts to solving Problem OPT(s_X), where s_X is the worst-case scenario for X taking the upper bounds of the uncertainty intervals on the edges of X and the lower bounds on the rest, the complexity of the algorithm is 2^d times the complexity of Problem OPT. We obtain the following.

Theorem 4. *If the number d of nondegenerate intervals of uncertainty is fixed, then Problem ROB can be solved with the same order of complexity as Problem OPT (up to a multiplicative constant). If d is not fixed but is bounded by the logarithm of a polynomial function of m, and if Problem OPT is polynomially solvable, then Problem ROB can be solved in polynomial time as well.*

Remark. The result of Theorem 4 is due to the special structure created by interval data representation of uncertainty (and not only due to "small" number of extreme scenarios); to illustrate this, we note that in the discrete-scenario case Problems ROB.TREE and ROB.PATH are NP-hard even if there are only two possible scenarios [5].

Acknowledgements

This research was supported by a grant from the Natural Sciences and Engineering Research Council of Canada (NSERC) to the first author.

References

[1] R. Ahuja, T. Magnanti, J. Orlin, Network Flows: Theory, Algorithms, and Applications, Prentice Hall, Upper Saddle River, NJ, 1993.
[2] I. Averbakh, Minmax regret solutions for minimax optimization problems with uncertainty, Oper. Res. Lett. 27 (2) (2000) 57–65.
[3] I. Averbakh, On the complexity of a class of combinatorial optimization problems with uncertainty, Math. Programming 90 (2001) 263–272.

[4] M. Garey, D. Johnson, Computers and Intractability: A Guide to the Theory of NP-Completeness, Freeman, New York, 1979.

[5] P. Kouvelis, G. Yu, Robust discrete optimization and its applications, Kluwer Academic Publishers, Boston, 1997.

[6] R. Möhring, Scheduling under uncertainty: optimizing against a randomizing adversary. Proceedings of the Third International Workshop on Approximation Algorithms for Combinatorial Optimization Problems, Springer Lecture Notes in Computer Science, Vol. 1913, Springer, Berlin, 2000, pp. 15–26.

[7] R. Montemanni, L.M. Gambardella, A branch-and-bound algorithm for the robust spanning tree problem with interval data, Presented at the Conference of the European Chapter on Combinatorial Optimization (ECCO), Lugano, Switzerland, May 30–June 1, 2002.

[8] J.M. Mulvey, R.J. Vanderbei, S.A. Zenious, Robust optimization of large scale systems, Oper. Res. 43 (2) (1995) 264–281.

[9] H. Yaman, O.E. Karasan, M. Pinar, Longest path problem with interval data, Technical Report 9908, Department of Industrial Engineering, Bilkent University, 06533 Bilkent, Ankara, Turkey, 1999.

[10] H. Yaman, O.E. Karasan, M. Pinar, The robust spanning tree problem with interval data, Technical Report 9909, Department of Industrial Engineering, Bilkent University, 06533 Bilkent, Ankara, Turkey, 1999.