The Minmax Regret Shortest Path Problem with Interval Arc Lengths

Jun-Gyu Kang

Department of Industrial & Management Engineering,
Sungkyul University, Korea
jun-gyu.kang@sungkyul.ac.kr

Abstract

This paper considers the shortest path problem on directed acyclic graphs, where the uncertainty of input data (lengths of arcs) is modeled in the form of intervals. In order to handle the interval data the minmax criterion to the regret values is applied, where the original objective function with interval coefficients is transformed into that of finding the least maximum worst-case regret, which is called the minmax regret criterion. A heuristic algorithm exploiting properties to reduce the computational times and to improve the solution quality is developed and its performance is shown with computational experiments.

Keywords: Shortest Path, Interval Data, Minmax regret

1. Introduction

The classical shortest path problem on directed acyclic graphs (DAGs) is a very well-known problem with lots of applications, i.e., route planning for transportation networks or for telecommunication networks, and community clustering for E-commerce [1]. Its deterministic version can be solved efficiently, for example, by using Dijkstra's algorithm or using Bellman-Ford’s algorithm. For large scale problems, meta-heuristics, i.e., the ant-colony optimization [2] or the genetic algorithm [3], can be applied. But in practical applications, the exact values of input data like lengths of arcs or duration of activities are often unknown in advance, due to the uncertain travel times on transportation networks and the uncertain transmission delays on telecommunication or on electric power networks [4]. In such cases, the most common approach to determine the optimal path is that the uncertainty is characterized by random distributions and the expected value of a utility function is maximized [5]. Nevertheless, the stochastic approach has limits to its application since it is often hard to determine proper random distributions due to the variable nature of the world or to a lack of information about a considered system.

With the least assumption about ambiguous variables, the simplest form of the uncertainty representation to adopt is that the value of arc lengths lies within a given range, independently from the values taken by the other parameters [6]. Here, the possible realization of the arc length corresponds to the interval ranges defined by known lower and upper bounds, but the actual realization and the probability distribution function are unknown. For the shortest path problem with interval data, the minmax regret criterion can be applied to find the relatively robust path [7]. Here, the regret denotes the deviation of length between a given path and the optimal path under a certain realization of arc lengths. Note that it is also called the relative robustness criterion or the robust deviation criterion.
The Minmax Regret Shortest Path Problem with interval arc lengths (MRSPP hereafter) is identifying the one; of which maximum regret over all possible realization of arc lengths is minimized, and a mixed integer programming formulation of the problem is presented in [8]. It is hard to obtain the exact solution, since the MRSPP is $NP$-hard even if the network is directed, acyclic, has a layered structure, and all bounds of intervals of uncertainty belong to $\{0, 1\}$ [9, 10]. As a practical approach for large scale problems, an enumerative greedy algorithm for the minmax regret longest path problem, where the $k$ non-dominated paths are generated by a branch-and-bound algorithm, was developed in [11]. Here, the non-dominated path indicates the path that is the shortest in at least one realization of arc lengths. Note that the longest path problem problem on DAGs can be transformed to the shortest path problem on DAGs, without loss of generality. Similar work was done by [12], which enumerates the $k$ shortest paths with improvement rules. In both heuristic algorithms, generating the paths until the near-optimal solution comes, takes extensive computational time, as checking the non-dominancy is $NP$-complete and generating the $k$ non-dominated paths is $NP$-hard [13]. As a constructive method, heuristic algorithms to estimate the maximum regret of the path from the source node to each intermediate node were developed in [14, 15].

This paper expands results in [14] by refining some properties, which reduces computational times and improves solution quality. It works for general directed graphs, although the performance of the suggested algorithm is illustrated with directed acyclic, layered graphs.

This paper is organized as follows: In Section 2, we describe the problem as that of finding the shortest path in a directed acyclic graph with interval arc lengths. In Section 3, some properties and an algorithm by exploiting the properties are suggested. In Section 4, the computational experiments on directed acyclic, layered graphs with small widths are presented and the results compared with those of the algorithm in [15] are followed. Finally, conclusions can be found in Section 5.

2. Problem Description and Preliminaries.

Let $G = (N, A)$ denote a directed graph, where $N$ is the set of nodes and $A$ is the set of arcs. An interval $[l_{ij}, u_{ij}]$ associated with each arc $(i, j) \in A$ represents the range of possible length for each arc, and is independent from the value of other arcs. Then, the shortest path problem can be formulated as a traditional linear programming model as follows:

$$(P) \quad \text{Min} \quad \sum_{(i, j) \in A} c_{ij} x_{ij} \quad (1)$$

Subject to

$$\sum_{[j \mid (i, j) \in A]} x_{ij} - \sum_{[j \mid (j, i) \in A]} x_{ji} = \begin{cases} 1 & \text{if } i = o, \\ -1 & \text{if } i = t, \\ 0 & \text{otherwise}, \end{cases} \quad (2)$$

$$x_{ij} \in \{0, 1\}, \quad \forall (i, j) \in A, \quad (3)$$

where $l_{ij} \leq c_{ij} \leq u_{ij}$ for all $(i, j) \in A$, and $o$ and $t$ denote the origin node and the terminal node, respectively. $x_{ij}$ is a binary decision variable, where $x_{ij} = 1$ if arc $(i, j)$ belongs to the shortest path, $x_{ij} = 0$ otherwise. In case where all of $c_{ij}$ are deterministic, the
problem can be solved by using the well-known Dijkstra’s algorithm and other efficient algorithms. However, in this paper, since $c_{ij}$ is assumed to be given in the form of interval $[l_{ij}, u_{ij}]$, no path can always be the shortest path over all realization of arc lengths and thus, the minmax regret criterion is applied as a surrogate measure. The MRSPP can be formally described as follows.

A particular realization of arc’s length, such that $c_{ij}^s \in [l_{ij}, u_{ij}]$ is fixed $\forall (i, j) \in A$, is called a scenario $s$. Thus, every scenario represents a specific realization of arc’s lengths, which possibly occurs with a positive, but unknown probability. Suppose a given arbitrary scenario $s$, path $x$ and path $y$ from $o$ to $t$, the regret of path $x$ against path $y$ is defined by the difference between the lengths of paths as follows:

$$ R(x, y, s) = |c^x - c^y| $$  \hspace{1cm} (4)

Let $\Phi$ be the set of all possible paths from $o$ to $t$ and $\Gamma$ be the set of all scenarios. Then the maximal regret of a given path $x$ can be defined as follows:

$$ R_{\max}(x) = \max_{s \in \Phi, y \in \Gamma} c^x - c^y $$  \hspace{1cm} (5)

In equation (5), the scenario $s$ and the path $y$ that maximize the right hand side is called the worst-case scenario and the worst-case alternative for $x$, respectively. Equation (5) can be solved straightforwardly by setting arc’s length, which is called $\epsilon$-consistency by [16], as follows:

$$ c_{ij} = \begin{cases} u_{ij} & \text{if arc}(i, j) \in x \\ l_{ij} & \text{otherwise} \end{cases} $$  \hspace{1cm} (6)

For a given path $x$, the arc’s length can be fixed as equation (6) to maximize equation (5) and the scenario is called the worst-case scenario for $x$. Equation (6) signifies that it is sufficient to deliberate the extreme point of $c_{ij} \in [l_{ij}, u_{ij}]$ while computing the maximum regret, and thus the number of possible scenarios is $|S| = 2^m (|A| = m)$. Under the worst case scenario $s$ is induced by the given path $x$, the shortest path maximizes the right-hand side of equation (5) and thus it is called the worst-case alternative for $x$ or the worst-case alternative induced by the scenario $s$ [12]. So, using equation (6), equation (4) and (5) can be simply written as follows:

$$ R_{\max}(x) = R(x, y) = c^x - c^y $$  \hspace{1cm} (7)

where, $c^x$ and $c^y$ denote the cost vector induced by $\epsilon$-consistency and $y$ denotes the worst-case alternative of $x$, in that, the shortest path derived from $c^x$ and $c^y$.

Then the MRSPP can be re-defined as follows:

$$ \min_{x \in \Phi} R_{\max}(x) = \min_{x \in \Phi, y \in \Gamma} \max_{s \in \Phi} (c^x - c^y) $$  \hspace{1cm} (8)

Applying equation (6), a mixed integer programming formulation is presented in [8] as follows:

(MRSP) Min $\sum_{(i, j) \in A} u_{ij}x_{ij} - y_i$  \hspace{1cm} (9)
\[
\begin{align*}
\text{subject to} & \quad y_j \leq y_i + l_{ij} + (u_{ij} - l_{ij})x_{ij} \quad \forall (i, j) \in A, \\
& \quad y_j \geq 0 \quad \forall i \in N,
\end{align*}
\]

where \( x_{ij} = 1 \), if arc\((i, j)\) belongs to the minmax regret shortest path; \( x_{ij} = 0 \) otherwise. \( y_j \) holds the cost of the shortest path from \( o \) to node \( i \) in the scenario induced by the path defined by \( x \) variables. The significant inequalities of the formulation are those in (10), which determine the length of each arc \( c_{ij} = \{l_{ij}, u_{ij}\} \) to keep \( c\)-consistency induced by \( x \) variables. The remaining constraints are basically the same as those of the classical shortest path problem described in equation (2) and (3).

3. The Algorithm

In this section, we present properties, which reduce computational times and improve the quality of the solution, develop rules based on the properties and build an efficient heuristic algorithm exploiting the rules.

3.1 Main Ideas

First, we suggest a property to reduce the computation time to obtain the worst-case alternative path for a given solution path. From the minmax regret and the \( c\)-consistency defined earlier, it is easy to find the worst-case alternative for a given solution as follows:

**Property 1** [16]. The worst-case alternative for a given solution \( x \) can be determined as the shortest path from the origin \( o \) to the terminal \( t \) in the scenario where each arc on \( x \) has cost \( c_{ij} = u_{ij} \) and the rest has cost \( c_{ij} = l_{ij} \).

Property 1 is already described in Section 2, through equation (5), (6) and (7). From Property 1, the following rule can be derived directly as follows:

**Rule 1: Initiation.** For a null path from \( o \) to \( o \), the scenario \( r \) which maximizes the regret is that \( c_{ij} = l_{ij} \) for all arcs and the worst case alternative is the shortest path where \( c_{ij} = l_{ij} \) for all arcs.

From Rule 1, the algorithm initiates a null path with the corresponding worst-case alternative and the maximum regret. Then, if a partial path from \( o \) to \( i \in N \) is determined, the worst-case alternative and the maximum regret can be found easily by extending Property 1. Given a partial path \( x \) from \( o \) to \( i \in N \), the worst case scenario \( s \) is that \( c_{ij} = u_{ij} \) for each arc \((i, j)\) on \( x \) and \( c_{ij} = l_{ij} \) for the rest, and the worst case alternative is the shortest path from \( o \) to \( t \) under the scenario \( s \).

However, to avoid solving the shortest path problem at each node, the following proposition can be used by expanding the partial path to an immediate successor arc. In the proposition given below, \( x^i \) denotes the minmax regret partial path to \( i \), defined as the partial path from \( o \) to \( i \in N \), which has the least maximum regret among all feasible partial paths from \( o \) to \( i \) over all possible scenarios. And, \( y^i \) denotes the worst-case
alternative path from $o$ to $t$ for $x^i$. Suppose that $x^i$ and $y^i$ are given for node $i \in N (i \neq t)$, let node $j \in N$ be an immediate successor of node $i$ and $x_{ij}^i$ be $x^i \cup \text{arc}(i,j)$.

**Proposition 1.** Given a minmax regret partial path $x^i$ and the corresponding worst-case alternative $y^i$, if $\text{arc}(i,j) \notin y^i$ then the path $x_{ij}^i$ is a minmax regret partial path from $o$ to $\text{arc}(i,j)$ and $y^i$ is the corresponding worst-case alternative.

**Proof.** Suppose that $\text{arc}(i,j)$ is not included in any worst-case alternative for path $x^i$. Then, by the definition of the minmax regret shortest path derived from equation (8), we have,

$$\min_{x \in \Lambda} R_{\max}(x) = \min_{x \in \Lambda} \left( u_{ij} - R_{\max}(x) \right) = u_{ij} - \min_{x \in \Lambda} R_{\max}(x),$$

where, $\Lambda$ denotes the set of all feasible paths from $o$ to node $i$. From the assumption of Proposition 1, we have,

$$\min_{x \in \Lambda} R_{\max}(x) = R(x^i,y^i).$$

Using equation (7), the maximum regret of the partial path $x_{ij}^i$ can be calculated as follows:

$$R_{\max}(x_{ij}^i) = R(x_{ij}^i,y^i) = u_{ij} - R(x^i,y^i),$$

From (12), (13) and (14),

$$\min_{x \in \Lambda \cup \text{arc}(i,j)} R_{\max}(x) = u_{ij} - R(x^i,y^i) = R_{\max}(x_{ij}^i),$$

This completes the proof.

Proposition 1 allows $\text{arc}(i,j)$ to take over the minmax regret partial path and the worst-case alternative from its predecessor, if the condition of Proposition 1 is satisfied. Therefore, we can avoid searching all feasible paths to find the minmax regret partial path from $o$ to $i \in N, i \neq t$. The path expanding rule, derived from Proposition 1, can be summarized as follows:

**Rule 2: Expanding a Path.** Given a minmax regret partial path $x^i$, it can be easily extended to $x_{ij}^i$ with a supplemented $\text{arc}(i,j)$, if $\text{arc}(i,j)$ is not on $y^i$.

If Proposition 1 is not satisfied, in that $\text{arc}(i,j)$ is on $y^i$, the worst-case alternative, $y_{ij}^i$, for $x_{ij}^i$ can be recomputed by Property 1. But, it is not guaranteed that $x_{ij}^i$ is a minmax regret partial path and has the least maximum regret among all candidate paths, of which worst-case alternative is $y_{ij}^i$. The following simple rule can be used to find a better candidate path for a given alternative $y_{ij}^i$. 

---

**Copyright © 2013 SERSC**
Rule 3: Generating a Candidate Path for a Given Alternative Path. Given candidate \( x^i_{+ij} \) and alternative \( y^i_{+ij} \), let \( G' \) be a sub graph of \( G \), where arcs on \( y^i_{+ij} \) but not on \( x^i_{+ij} \) are uninvolved. Then, a new candidate path, \( x^i_{-ij} \), is the shortest path from \( o \) to \( \text{arc}(i,j) \), where \( c_{ij} = u_{ij} \) for all arcs on \( G' \).

The Rule 3 can be derived from the following Theorem 1.

**Lemma 1.** \( y^i_{+ij} \) is the worst-case alternative for \( x^i_{+ij} \) on \( G \).

**Proof.** Let \( s^1 \) be the scenario, where \( c_{ij} = u_{ij} \) for all arcs on \( x^i_{+ij} \) and \( c_{ij} = l_{ij} \) for the rest on \( G \), and \( s^2 \) be the scenario, where \( c_{ij} = u_{ij} \) for all arcs on \( G' \) and \( c_{ij} = l_{ij} \) for the rest on \( G \). The cost of each arc under \( s^2 \) is greater than or equal to that under \( s^1 \), since more arcs are set to \( u_{ij} \) under \( s^2 \) and the rest are the same. Thus, the shortest path under \( s^1 \) is shorter than or equal to the shortest path under \( s^2 \). The cost of arc on \( y^i_{+ij} \) is the same under \( s^1 \) and \( s^2 \), since the arcs on \( y^i_{+ij} \) are pegged for \( G' \) or the costs are already set to \( u_{ij} \). The shortest path on \( G \) under \( s^2 \) is the worst-case alternative for each path on \( G \) or \( G' \) under \( s^2 \). By definition, \( y^i_{+ij} \) is the shortest path on \( G \) under \( s^1 \). Therefore, \( y^i_{+ij} \) is the shortest path on \( G \) under \( s^2 \). Since \( x^i_{-ij} \) is a path on \( G' \) under \( s^2 \), \( y^i_{+ij} \) is the worst-case alternative for \( x^i_{+ij} \). The result follows.

From Lemma 1, the following Theorem 1 can be derived.

**Theorem 1.** The maximum regret of \( x^i_{-ij} \) is less or equal to the maximum regret of \( x^i_{+ij} \).

**Proof.** From Lemma 1, it is obvious that \( y^i_{+ij} \) is the worst-case alternative for \( x^i_{-ij} \) and \( x^i_{+ij} \). By definition of \( G' \), \( x^i_{+ij} \subset G' \) and \( x^i_{-ij} \subset G' \). By definition of \( x^i_{-ij} \) in Rule 4, \( x^i_{-ij} \) is the shortest path under \( s^2 \), which embraces the scenario induced by \( x^i_{+ij} \). By definition of the maximum regret described in equation (7), \( c^i_y \) is constant since \( y^i_{+ij} \) is given. \( c^o x^i_{-ij} \) is less than or equal to \( c^o x^i_{+ij} \). This completes the proof.

Note that, it is normally not guaranteed that there always exists a minmax regret shortest path which comprises an arbitrary arc\((i,j)\), since some arcs, which are called non-weak arcs, are never in an optimal solution of the MRSPP [8]. In this paper, it is supposed that the given graph is exclusive of non-weak arcs.
3.2 Pseudo-code

A heuristic algorithm exploiting the properties been described in Section 3.1 is proposed with the pseudo-code briefed in Figure 1. The algorithm starts at the origin node by initializing the maximum regret, \( r_j \), for each node \( j \) at line 0-1 in Figure 1. And then the null path and the corresponding worst-case alternative are determined by Rule 1 at line 2. Then, an iterative procedure is performed to find the minmax regret partial path by examining each incoming arc \((i, j)\) for each node \( j \) at line 3-15. At each step, if \( \text{arc}(i, j) \not\in y^i \), at line 5, Rule 2 is applied to get the worst-case alternative and the corresponding maximum regret, straightforwardly, at line 6-7. Else, Rule 3 can be applied to find a better candidate path from \( o \) to \( \text{arc}(i, j) \) at line 9-10. When all incoming arcs to node \( j \) are probed, the least maximum deviated path among them is taken as the minmax regret partial path for node \( j \) at line 12-13. This iterative procedure is repeated onto node \( t \).

![Figure 1. Pseudo-code of the Proposed Algorithm](image)

4. Computational Experiments

This section presents some computational results to evaluate the performance of the proposed algorithm RDSP, described in the previous section. To demonstrate the performance and the efficiency, MILP model in [8] was adopted to find the optimal solution and the algorithm in [15] was selected to show the accuracy of the proposed algorithm.

4.1 Network Model

The algorithm has been tested on graphs suggested in [8], called Karaşan’s graphs. As stated in [8], these graphs are suitable to simulate telecommunication networks. Karaşan’s graphs are directed acyclic and layered graphs with a small width. Note that a directed acyclic graph is a directed graph with no directed cycles, and a layered graph
is a connected graph where layers partition the nodes. Each edge, which has a nonnegative integral weight, connects only nodes in successive layers. The width is the greatest number of vertices in any layer. Particularly, in Karaşan’s graphs, the width of each layer is the same except for the first and the last layer, which are the origin and the terminal, respectively, and each node is connected to every node in the successive layer.

As stated in [17], each interval cost \([l_{ij}, u_{ij}]\) is obtained by generating a random number \(c_{ij} \in [1, c]\) and by randomly selecting \(l_{ij}\) in \([(1-d)c_{ij}, (1+d)c_{ij}]\) and \(u_{ij}\) in \([l_{ij}, (1+d)c_{ij}]\), where \(0 < d < 1\). Here, parameter \(c\) and \(d\) are used to control the degree of uncertainty. The bigger \(c\) spreads the center of interval for each arc, the bigger \(d\) widens the interval.

### 4.2 Experimental Result

The algorithm is coded in C++ using the Standard Template Library. All the tests were done on a personal computer with Intel Core2Duo @ 2.33GHz CPU / 1GB main memory. To obtain the optimal solution based on Karaşan’s MILP model, Cbc (Coin-or branch and cut, open-source MILP solver [18]) ver. 1.1.2 was applied. Although it is said that Cbc is slower than commercial MILP solvers, \(i.e., 5\) times slower than ILOG CPLEX 12.0 [19], the result of Cbc is only used for calculating the relative errors of the proposed algorithm from the optimal solution, not for comparing the computational time.

Tests were done with combinations of 4 parameters, \(i.e.,\) the number of layers\(\in\{50, 100\}\), width\(\in\{2, 4\}\), \(c\in\{10, 20\}\), and \(d\in\{0.3, 0.9\}\). For each combination of parameters, the test was repeated 100 times with randomly generated instances.

The computational results are summarized in Table 1. In the first part of the table shows the computational time of Cbc and the proposed algorithm. Note that ‘mean’ indicates the arithmetic mean, STD indicates the standard deviation. The computational time of the heuristic algorithm in [15] is almost 0 and so not presented. Although the computational time for the exact solution increases about 50 times as the width of networks increases from 2 to 4, the computational time of the proposed algorithm increase only twice as shown in the second column of Table 1. The second part shows the relative errors between the algorithm results and the exact solution. On average, the relative errors of the proposed algorithm are much less than that of the fast heuristic algorithm in [15]. Especially, it shows that the performance of the algorithm is robust against any combination of the parameters, since the relative errors are stable around 1%. And the third part ‘Correct Ratio (%)’ shows the number of problems, that the algorithms optimally solved, out of 100 instances. It shows that the proposed algorithm can find optimal solutions with relatively high percentages. The computational results show that the proposed algorithm outperforms other heuristic algorithms developed previously, to author’s knowledge, with the most accurate solutions.
5. Conclusions

This paper considered a generalization of the shortest path problem on DAGs, in which the lengths of arcs are given in the form of intervals. To treat the ambiguity of the arc lengths, we applied the minmax regret criterion and transformed the shortest path problem with interval data into the minmax regret shortest path problem. The solution algorithm includes 3 major rules based on properties proved in this paper, which ease the complexity to find the minmax regret partial path and the corresponding worst-case alternative. The computational experiments were carried out, and the results show that the proposed algorithm provides optimal solutions with relatively high percentage or near optimal solutions with about 1% of relative errors from optimal solutions, within a reasonable amount of computation time.

Acknowledgements

This research was supported by the research fund of Sungkyul University under grant No. 2011-1142-1.

References


Authors

Jun-Gyu KANG

He is currently an Assistant Professor in Department of Industrial and Management Engineering at Sungkyul University, Republic of Korea. He received a B.S. degree in Industrial Engineering from Busan National University in 1998 and the M.S. in Industrial Engineering from Korea Advanced Institute of Science and Technology in 2000. He did his PhD in Mechanical Engineering at the Swiss Federal Institute of Technology – Lausanne (École Polytechnique Fédéral de Lausanne) in 2005. From 2006 to 2007, he worked at Grenoble Institute of Technology (Institut Polytechnique de Grenoble, France) as a research fellow. His areas of research include Robust Optimization, Product Lifecycle Management, Product Design and Development for End-of-Life, and System Modeling and Simulation. E-mail: jun-gyu.kang@sungkyul.ac.kr.