Spanning Tree Problem of Uncertain Network

Jin Peng
Institute of Uncertain Systems
Huanggang Normal University
Hubei 438000, China
Email: pengjin01@tsinghua.org.cn

Shengguo Li
College of Mathematics & Computer Science
Huanggang Normal University
Hubei 438000, China
Email: lisg@hgnu.edu.cn

Abstract—In this paper we consider a spanning tree problem of uncertain network, which is a natural uncertain variation of the deterministic minimum spanning tree problem. By means of uncertainty theory, we propose three types of minimum spanning tree models for uncertain network, i.e., expected minimum spanning tree, α-minimum spanning tree, and distribution minimum spanning tree. Generally, we can theoretically give the inverse uncertainty distributions of uncertain spanning trees of an uncertain network. A 99-table algorithm for finding the inverse uncertainty distribution of uncertain spanning tree is given. Some numerical examples are illustrated according to the graphical topology structure of the uncertain spanning tree problem.

Keywords: network optimization, expected minimum spanning tree, α-minimum spanning tree, distribution minimum spanning tree, inverse uncertainty distribution, 99-table algorithm

I. INTRODUCTION

The minimum spanning tree (MST) problem [7] is one of the most important network optimization problems, which has widely and important applications in many fields such as transportation, communications, logistics, etc.

One example of MST problem would be a TV cable company laying cable to a new neighborhood. If it is constrained to bury the cable only along certain paths, then there would be a network in which points are connected by those paths. Some of those paths might be more expensive because of longer or being buried deeper, and these paths would be represented by edges with larger weights. A spanning tree for that network would be a subset of those paths that have no cycles but still connects to every house. There might be several spanning trees possible. A minimum spanning tree would be one with the lowest total cost.

The deterministic MST problem has been well studied and many efficient algorithms have been developed by many researchers [6], [14], [17], [20]. There are now two algorithms commonly used, Kruskal’s algorithm [10] and Prim’s algorithm [16]. For a certain network, if the edge weights of the network are fixed, then we can calculate the minimum spanning tree by the algorithms as mentioned above.

However, network optimization in real world is usually made in uncertain environment and various types of uncertainty are frequently encountered in practice. For example, usually we are not sure the weight of some edges of a new network. A network is called uncertain if parts or all of the edge weights are uncertain. For an uncertain network, that is, the edge weights of the network are uncertain, some researchers regard the edge weights as random variables, and they use probability theory [1], [4] to study the probabilistic minimum spanning tree (PMST) problem. It may be assumed that not all the points are deterministically present, but are present with certain probability [18]. Meanwhile, some researchers regard the edge weights as fuzzy variables, and they use fuzzy theory [5], [8] to study the fuzzy minimum spanning tree (FMST) problem. More complicatedly, some researchers regard the edge weights as fuzzy random variables [9], and they use fuzzy random theory to study the fuzzy random minimum spanning tree (FRMST) problem. Differently, some researchers regard the edge weights as interval data [3], [19], and they use interval analysis theory to study the interval data minimum spanning tree (IDMST) problem. Uncertain spanning tree (UST) problem is a natural uncertain variation of the deterministic MST, in which not all the edge weights are deterministically present, but are present with uncertain variables.

In order to deal with some uncertain phenomena, uncertainty theory was founded by Liu [11] in 2007 refined by Liu [13] in 2010, and became a branch of mathematics based on the normality, self-duality, countable subadditivity, and product measure axioms. For exploring the recent developments of uncertainty theory, the readers may consult [12].

How to define and find the minimum spanning tree of a connected uncertain network? For an uncertain network, the weights of edges are approximately estimated by the expert. In uncertainty theory, we can obtain expert’s experimental data by uncertain statistic and give the uncertainty distribution [13]. Since the weights of the spanning trees of an uncertain network are all uncertain variable, there are no ranking relationship among them. Our key task is to provide some types of minimum spanning tree for uncertain network, e.g., expected minimum spanning tree, α-minimum spanning tree, and distribution minimum spanning tree. Another way is to give theoretically the uncertainty distribution or inverse uncertainty distribution of uncertain spanning tree of an uncertain network by taking into account its graphical topology structure. Thanks to 99-table algorithm for finding the inverse uncertainty distribution of uncertain variable proposed by Liu [13], we can represent the wanted inverse uncertainty distribution by simulation.

The paper is organized as follows: Section 2 presents some preliminary concepts and results selected from uncertainty theory, the readers may consult [12].
theory. In Section 3, the UST problem is introduced and some properties are investigated. In Section 4, three types of minimum spanning tree models for UST problem are presented, which includes expected minimum spanning tree, minimum spanning tree and distribution minimum spanning tree. A 99-

able algorithm for finding the inverse uncertainty distribution of UST in uncertain network is exhibited in Section 5. Section 6 illustrates some examples. The last section contains the conclusions.

II. PRELIMINARIES

Uncertainty theory provides an efficient tool to solve UST problem. In this section, we present some preliminaries from uncertainty theory.

Let be a nonempty set, and a -algebra over . For any , Liu [11] presented an axiomatic uncertain measure to express the chance that uncertain event occurs. The set function satisfies the following three axioms: (i) (Normality) ; (ii) (Self-Duality) for any ; (iii) (Countable Subadditivity) For every countable sequence of events , we have . The triplet is called an uncertainty space and an uncertain variable is defined as a measurable function from this space to the set of real numbers (Liu [11]).

An uncertain variable can be characterized by its uncertainty distribution : , which is defined by Liu [11] as follows

\[ \Phi(x) = M\{\gamma \in \Gamma \mid \xi(\gamma) \leq x\}. \] (1)

Peng and Iwamura [15] have proved that a function : , is uncertainty distribution if and only if it is an increasing function except and .

Let be an uncertain variable with uncertainty distribution . Then the inverse function is called the inverse uncertainty distribution of .

The expected value of uncertain variable is defined by Liu [11] as

\[ E[\xi] = \int_0^{+\infty} M\{\xi \geq r\}dr - \int_{-\infty}^0 M\{\xi \leq r\}dr \] (2)

provided that at least one of the two integrals is finite.

As a useful representation of expected value, it has been proved by Liu [11] that

\[ E[\xi] = \int_0^1 \Phi^{-1}(\alpha)d\alpha \] (3)

where is the inverse uncertainty distribution of uncertain variable .

Liu [11] introduced the independence concept of uncertain variables. The uncertain variables are independent if and only if

\[ M\left\{ \bigcap_{i=1}^m \{\xi_i \in B_i\} \right\} = \min_{1 \leq i \leq m} M\{\xi_i \in B_i\} \] (4)

for any Borel sets , of .

A real-valued function is said to be strictly increasing if

\[ f(x_1, x_2, \cdots, x_n) \leq f(y_1, y_2, \cdots, y_n) \] (5)

whenever for and

\[ f(x_1, x_2, \cdots, x_n) < f(y_1, y_2, \cdots, y_n) \] (6)

whenever for .

Liu [11] introduced the following useful theorem to determine the distribution function of the strictly increasing function of uncertain variables.

**Theorem 2.1:** Let , be independent uncertain variables with uncertainty distributions , , respectively. If is a strictly increasing function, then is an uncertain variable with inverse uncertainty distribution

\[ \Phi^{-1}(\alpha) = f(\Phi_1^{-1}(\alpha), \Phi_2^{-1}(\alpha), \cdots, \Phi_n^{-1}(\alpha)). \]

III. UNCERTAIN SPANNING TREE PROBLEM

In mathematics and computer science, graph theory is the study of graphs, mathematical structures used to model pairwise relations between objects from a certain collection. We now refer to some elementary definitions from graph theory [2].

A. Deterministic Minimum Spanning Tree Problem

Given a connected weighted undirected graph , a spanning tree of that graph is a subgraph , such that , which is a tree and connects all the vertices together. A single graph can have many different spanning trees. A minimum spanning tree is then a spanning tree with weight less than or equal to the weight of every other spanning tree. A minimum spanning tree contains a subset of the edges that forms a tree and includes every vertex, where the total weight of all the edges in the tree is minimized.

Many efficient algorithms for deterministic MST problem have been developed by many researchers. The first algorithm for finding a minimum spanning tree was developed by Borůvka in 1926. There are now two algorithms commonly used, Kruskal’s algorithm and Prim’s algorithm. Prim’s algorithm was developed in 1930 by Czech mathematician Jarník and later independently by computer scientist Prim in 1957 and rediscovered by Dijkstra in 1959. Therefore it is also sometimes called the DJP algorithm, the Jarník algorithm, or the Prim-Jarník algorithm.

**Algorithm 3.1:** (Kruskal’s Algorithm)

- Create a forest (a set of trees), where each vertex in the graph is a separate tree;
- Create a set containing all the edges in the graph;
- While is nonempty and is not yet spanning
  - Remove an edge with minimum weight from ;
  - If that edge connects two different trees, then add it to the forest, combining two trees into a single tree;
  - Otherwise discard that edge.
Algorithm 3.2: (Prim’s Algorithm)

- Input: A non-empty connected weighted graph with vertices $V$ and edges $E$;
- Initialize: $V_{\text{new}} = x$, where $x$ is an arbitrary node (starting point) from $V$, $E_{\text{new}} = $;
- Repeat until $V_{\text{new}} = V$:
  - Choose an edge $(u, v)$ with minimal weight such that $u$ is in $V_{\text{new}}$ and $v$ is not (if there are multiple edges with the same weight, any of them may be picked);
  - Add $v$ to $V_{\text{new}}$, and $(u, v)$ to $E_{\text{new}}$;
- Output: $V_{\text{new}}$ and $E_{\text{new}}$ describe a minimal spanning tree.

B. Uncertain Spanning Tree Problem

In most scenarios, it is assumed that the edge weights are fixed, but this is not always true. For example, links in a communication network can malfunction or degrade as a result of congestion, accidents, weather, etc. More generally, the edges of a time-varying network can assume several states. Therefore, a deterministic network is not able to realistically model the characteristics of such networks, and so the network topology should be modelled by an uncertain network.

In an uncertain network, each spanning tree of it is called an uncertain spanning tree. Uncertain spanning tree (UST) problem is a natural uncertain variation of the deterministic spanning tree problem. It is assumed that the edge weights (costs) are not precisely known and they are specified as uncertain variables. Uncertainty theory is applied to characterize a spanning tree under uncertain costs. Since the weight of the spanning tree of an uncertain network is still an uncertain variable, our key task and the most important work is to give its uncertainty distribution or inverse uncertainty distribution.

Let $N$ be an uncertain network with edge weights of independent uncertain variables $\xi_1, \xi_2, \ldots, \xi_n$. The weight of the uncertain spanning tree may be comprehended as an uncertain variable $\xi = f(\xi_1, \xi_2, \ldots, \xi_n)$, where $f$ is a $n$-dimensional strictly increasing function. By Theorem 2.1, we have

**Theorem 3.1:** Let $N = (V, E)$ be an uncertain network with $n$ edges and the weights of independent uncertain variable $\xi_1, \xi_2, \ldots, \xi_n$. Then the weight of uncertain spanning tree of $G$ is an uncertain variable $\xi = f(\xi_1, \xi_2, \ldots, \xi_n)$ and its inverse uncertainty distribution is

$$\Phi^{-1}(\alpha) = f(\Phi_1^{-1}(\alpha), \Phi_2^{-1}(\alpha), \ldots, \Phi_n^{-1}(\alpha)).$$

**Theorem 3.2:** Let $N = (V, E)$ be an undirected network with $m$ vertices and $n$ edges and the uncertain weights represented by independent uncertain variable $\xi_i(i = 1, 2, \ldots, n)$. Assume that the $n = m - 1$ edges of the uncertain spanning tree are of uncertain weights $\xi_i(j = 1, 2, \ldots, m - 1)$. Then the inverse uncertainty distribution of the uncertain spanning tree of $T$ is

$$\Phi^{-1}(\alpha) = \Phi_{i_1}^{-1}(\alpha) + \Phi_{i_2}^{-1}(\alpha) + \cdots + \Phi_{i_{m-1}}^{-1}(\alpha).$$

This result implies that we can directly find the inverse uncertainty distribution of uncertain spanning tree by way of inverse uncertainty distribution of each edge in the tree.

For a special graph of tree uncertain network, we can easily obtain the inverse uncertainty distribution of uncertain spanning tree itself.

**Theorem 3.3:** Let $T = (V, E)$ be a tree network with $m$ vertices and $n = m - 1$ edges with independent uncertain weights $\xi_i(i = 1, 2, \ldots, m - 1)$, respectively. Then the inverse uncertainty distribution of $T$ is an uncertain variable $\xi = \xi_1 + \xi_2 + \cdots + \xi_{m-1}$ and its inverse uncertainty distribution is

$$\Phi^{-1}(\alpha) = \Phi_1^{-1}(\alpha) + \Phi_2^{-1}(\alpha) + \cdots + \Phi_{m-1}^{-1}(\alpha).$$

**Theorem 3.4:** Let $C = (V, E)$ be a circle uncertain network with $m$ vertices and $n$ edges with independent uncertain weight $\xi_i(i = 1, 2, \ldots, m)$, respectively. Then the weight of the uncertain minimum spanning tree of $C$ is an uncertain variable $\xi$ and its inverse uncertainty distribution is

$$\Phi^{-1}(\alpha) = \bigwedge_{i=1}^n [\Phi_1^{-1}(\alpha) + \cdots + \Phi_{i-1}^{-1}(\alpha) + \cdots + \Phi_{m-1}^{-1}(\alpha) + \Phi_m^{-1}(\alpha)].$$

**Theorem 3.5:** Let $N = (V, E)$ be an undirected network with $m$ vertices and $n$ edges of independent uncertain variable weights $\xi_i(i = 1, 2, \ldots, n)$. Assume that the $n = m - 1$ edges of the uncertain spanning tree are of uncertain weight $\xi_i(j = 1, 2, \ldots, m - 1)$. Then the expected value of uncertain weight $\xi$ of the spanning tree of $G$ is

$$E[\xi] = E[\xi_{i_1}] + E[\xi_{i_2}] + \cdots + E[\xi_{i_{m-1}}].$$

This result tells us that we can find the whole expected weight of uncertain spanning tree by replace of expected weight of each edge in the uncertain spanning tree.

IV. THREE TYPES OF UNCERTAIN MINIMUM SPANNING TREES

We try to define the minimum spanning tree for uncertain network in different ways. Here we give three types of uncertain minimum spanning trees, which includes expected minimum spanning tree, $\alpha$-minimum spanning tree, and distribution minimum spanning tree.

**Definition 4.1:** Let $N = (V, E)$ be an uncertain network with $n$ edges and the weight of $i$-th edge be an uncertain variable $\xi_i(i = 1, 2, \ldots, n)$. Denote $\xi = (\xi_1, \xi_2, \ldots, \xi_n)$.

An uncertain spanning tree $T_\alpha$ is called the expected minimum spanning tree (EMST) if

$$E[W(T_\alpha, \xi)] \leq E[W(T, \xi)]$$

holds for all minimum spanning tree $T$, where $W(T, \xi)$ stands for the weight of spanning tree $T$ and $E[W(T, \xi)]$ is called the weight of expected minimum spanning tree.

**Definition 4.2:** Let $N = (V, E)$ be an uncertain network with $n$ edges and the weight of $i$-th edge be an uncertain variable $\xi_i(i = 1, 2, \ldots, n)$. Denote $\xi = (\xi_1, \xi_2, \ldots, \xi_n)$. 
An uncertain spanning tree $T_e$ is called the $\alpha$-minimum spanning tree ($\alpha$-MST) if
\[
\min \{ \omega \mid M(W(T_e, \xi) \leq \omega) \geq \alpha \}
\leq \min \{ \omega \mid M(W(T, \xi) \leq \omega) \geq \alpha \}
\]
or equivalently,
\[
\Phi^{-1}_e(\alpha) \leq \Phi^{-1}_T(\alpha)
\]
holds for all minimum spanning tree $T$, where $\alpha$ is a predetermined confidence level.

**Definition 4.3:** Let $N = (V, E)$ be an uncertain network with $n$ edges and the weight of $i$-th edge be an uncertain variable $\xi_i$ ($i = 1, 2, \ldots, n$). Denote $\xi = (\xi_1, \xi_2, \ldots, \xi_n)$.

An uncertain spanning tree $T_e$ is called the distribution minimum spanning tree (DMST) if
\[
\Phi_e(x) \leq \Phi_T(x)
\]
holds for all minimum spanning tree $T$ and for any $x \in \mathbb{R}$.

**V. An Algorithm for Uncertain Spanning Tree**

Generally speaking, finding the deterministic minimum spanning tree in an uncertain network becomes considerably harder when the edge weights are of different types of uncertain variables. In what follows, we present a 99-table algorithm for the uncertain spanning tree problem based on the theoretical results mentioned in above section. The basic idea behind our algorithm is to determine the inverse uncertainty distribution of the uncertain spanning tree by 99-table.

**99-Method:** In [13], Liu suggested that an uncertain variable $\xi$ with uncertainty distribution $\Phi$ is represented by a 99-table

\[
\begin{array}{c|c|c|c}
0.01 & 0.02 & \cdots & 0.99 \\
\hline
x_1 & x_2 & \cdots & x_{99}
\end{array}
\]

where 0.01, 0.02, \ldots, 0.99 in the first row are the values of uncertainty distribution $\Phi$, and $x_1, x_2, \ldots, x_{99}$ in the second row are the corresponding values of inverse uncertainty distribution $\Phi^{-1}(0.01), \Phi^{-1}(0.02), \ldots, \Phi^{-1}(0.99)$. That is to say, the uncertainty distribution $\Phi$ of uncertain variable $\xi$ can be discretely and approximately replaced by
\[
\Psi(x) = \begin{cases} 
0, & \text{if } x \leq x_1 \\
\alpha_i + (\alpha_{i+1} - \alpha_i) \frac{x - x_i}{x_{i+1} - x_i}, & \text{if } x_i \leq x \leq x_{i+1} \\
1, & \text{if } x > x_{99},
\end{cases} 
\]

(1 \leq i \leq 99)

Essentially, the 99-table is a discrete representation of uncertainty distribution $\Phi$. Then for any strictly increasing function $f(x)$, the uncertain variable $f(\xi)$ has a 99-table

\[
\begin{array}{c|c|c|c}
0.01 & 0.02 & \cdots & 0.99 \\
\hline
f(x_1) & f(x_2) & \cdots & f(x_{99})
\end{array}
\]

The 99-method may be extended to the 999-method if a more precise result is needed.

Therefore, we can give the inverse uncertainty distribution of the uncertain weight of uncertain network by 99-table as follows:

\[
\begin{array}{c|c|c}
0.01 & \cdots & 0.99 \\
\hline
f(\Phi^{-1}_1(0.01), \ldots, \Phi^{-1}_n(0.01)) & \cdots & f(\Phi^{-1}_1(0.99), \ldots, \Phi^{-1}_n(0.99))
\end{array}
\]

**VI. Example**

Now we report results from computational experiments that demonstrate our method.

An uncertain variable $\xi = Z(a, b, c)$, where $a, b, c$ are real numbers with $a < b < c$, is called a zigzag uncertain variable if its uncertainty distribution is
\[
\Phi(x) = \begin{cases} 
0, & \text{if } x \leq a \\
\frac{x - a}{2(b - a)}, & \text{if } a \leq x \leq b \\
\frac{x + c - 2b}{2(c - b)}, & \text{if } b \leq x \leq c \\
1, & \text{if } x \geq c.
\end{cases}
\]

It is easily calculated that the inverse distribution of the zigzag uncertain variable can be analytically expressed as
\[
\Phi^{-1}(\alpha) = \begin{cases} 
a + 2(b - a)\alpha, & \text{if } \alpha \leq 0.5 \\
2b - c + 2(c - b)\alpha, & \text{if } \alpha > 0.5.
\end{cases}
\]

In addition,
\[
E[\xi] = \int_0^1 \Phi^{-1}(\alpha) d\alpha = \frac{a + 2b + c}{4}.
\]

**Example 6.1:** Let $N = (V, E)$ be an uncertain network with 5 vertices and 9 edges with the weights of zigzag uncertain variables $\xi_i (i = 1, 2, \ldots, 9)$ as shown in Fig. 1.

Following the deterministic MST algorithm and Theorem 3.5, we can find the weight of EMST is 9, which consists of edges $e_{12}, e_{25}, e_{53}, e_{54}$.

Fig. 1. Uncertain Network
Now let us take $\alpha = 0.9$. Following the deterministic MST algorithm and Theorem 3.2, we can find the weight of 0.9-MST consists of edges $e_{12}, e_{25}, e_{53}, e_{54}$ with

$$\Phi^{-1}(0.9) = 1.8 + 2.8 + 2.8 + 4.8 = 12.2.$$