Longest Path Problems in DAGs with Probabilistic Weights

Chao Yan & Tingting Xiao
chaoyan@hexineda.com

April 2024

1 Introduction

This document describes the problems to be solved, including the Longest Path Problem in Directed Acyclic Graphs (DAGs) with a focus on probabilistic weights, identifying the top-k longest paths, and heuristic approaches for solving these challenges if no exact solution exists.

2 Problem Definition

2.1 Longest Path Problem with Probabilistic Weights

We encounter a unique challenge in finding the longest path in a Directed Acyclic Graph (DAG) where each edge weight is represented as a probabilistic value. Specifically, these weights are characterized by Gaussian distributions, making the problem significantly more complex than its deterministic counterpart.

Problem Statement: Consider a DAG with a single source and sink, where the weight of each edge is a positive, random value w_i, approximated by a Gaussian distribution $N(\mu_i, \sigma_i)$. For a path comprising weights $w_1, w_2, \ldots, w_n$, the 3σ path distance is defined as:

$$d_{3\sigma}(\text{path}) = \sum_{i=1}^{n} \mu_i + 3\sqrt{\sum_{i=1}^{n} \sigma_i^2}.$$

Take Figure 1 as an example, what's the longest path over the four paths?

This probabilistic nature of edge weights introduces a significant challenge, deviating from traditional methods like Dijkstra's algorithm, which efficiently solves the problem with deterministic weights.

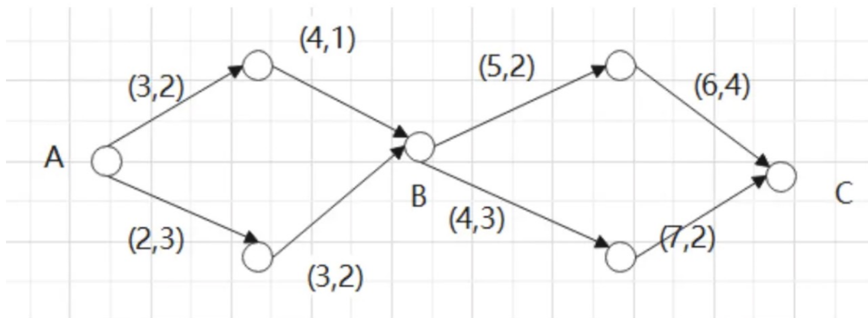

Figure 1: DAG with probabilistic weights.

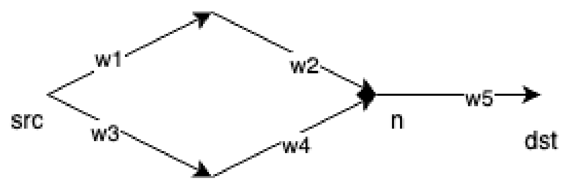

Figure 2: Local maximum may not be global maximum

Consider the example in Figure 2, let

$$
\begin{aligned}
\mu_1 + \mu_2 &= 1 \\
\sqrt{\sigma_1^2 + \sigma_2^2} &= 1.5 \\
\mu_3 + \mu_4 &= 2 \\
\sqrt{\sigma_3^2 + \sigma_4^2} &= 1 \\
\mu_5 &= 1 \\
\sigma_5 &= 2
\end{aligned}
$$

At node n, the path (w_1, w_2) is longer than (w_3, w_4) as $1+3\cdot 1.5 > 2+3\cdot 1$. However, by adding the edge $w5$, the path (w_3, w_4, w_5) is longer than (w_1, w_2, w_5) as $2+3\cdot\sqrt{2^2+1.5^2} = 9.5 < 3+3\cdot\sqrt{2^2+1} = 9.7$.

The problem is similar when adjusting 3σ path distance to σ path distance. And it should be same problem to find the shortest path with -3σ path distance.

2.2 Top-k Longest Paths Problem

Building upon the initial problem, we extend the challenge to finding the top-k longest paths within the DAG, considering the probabilistic nature of edge weights. This variation introduces further computational complexity and demands innovative solutions.

2.3 Heuristic Approach for the Top-k Longest Paths

If there does not exist efficient algorithm that finding the solution without enumerating all possible paths, heuristic algorithms are helpful as well. There are some related works, for example, an algorithm is proposed in [1] which records two partial paths with maximum μ and maximum $\mu + 3\sigma$. However, it doesn't ensure the longest path, nor provides the probability.

Thus, the third facet of our problem seeks a heuristic method capable of identifying the top-k longest paths with a specified probability (e.g. $> 99\%$) or providing a slightly larger set (e.g. 1.2k) containing the top-k paths.

We can assume σ is typically smaller than μ, e.g. $\sigma < 0.1\mu$.

3 Background and Motivation

The issue originates from static timing analysis (STA), which is tasked with calculating the delay along a circuit's path.

In the absence of Parametric On-Chip Variation (POCV), Graph-Based Analysis (GBA) is utilized to compute the delay across each timing arc and identify the longest path. Leveraging data derived from GBA, a Top-k algorithm is then employed to enumerate the top-k longest paths.

However, the introduction of POCV complicates the scenario significantly. While GBA still calculates the delay of timing arcs, it no longer can definitively pinpoint the longest path due to the variability introduced by POCV. Instead, an estimation approach is adopted, which involves merging branch distributions to approximate the delays, as illustrated in Figure 3.

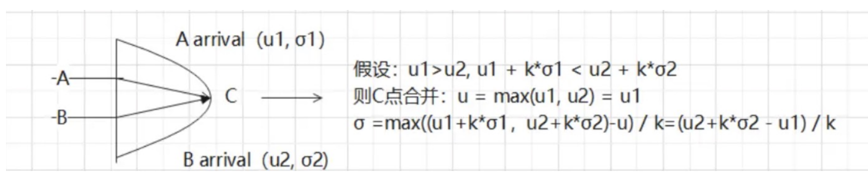

Figure 3: Merging branch distributions in GBA

4 Examples

Further examples and data pertaining to these problems are available upon request, aiming to foster deeper understanding and facilitate collaborative research efforts.

References

[1] Ahish Mysore Somashekar and Spyros Tragoudas. Efficient critical path selection under a probabilistic delay model. In *Proceedings of the on Great Lakes Symposium on VLSI 2017*, pages 185–190, 2017.