[Volume 3 issue 2 February 2015]
Page No.890-907 ISSN :2320-7167

INTERNATIONAL JOURNAL OF MATHEMATICS AND COMPUTER RESEARCH

# Stochastic Optimization for Routing Shortest Path in a Network

***Zahra M. Elsayed[1], Mustafa A. O.[2],M.E.Wahed[3]***
[1,2] Department of Mathematics,Faculty of Sience,Zagazig University
[3]Faculty of Computers and Information,Suez Canal University, EGYPT
E-mail: mewahed@yahoo.com

## ABSTRACT

In VLSI circuit design, graph algorithms are widely used and graph structure can model many problems. As technology continues to scale into nanometer design, the effects of process variation become more crucial and design parameters also change. Hence, taking stochastic variations into account, probability distributions are used as edge weights to form statistical graph structures. General applications in VLSI circuit design, such as timing analysis, buffer insertion, and maze routing, can be formulated as shortest path problems using a statistical graph model. The solution of any such graph problem will surely have a statistical distribution for its cost function value. The mean and variance, square of standard deviation, values are used as a pair of weight values on a graph to represent the stochastic distribution on each edge. For the stochastic shortest path problem, we observe that the objective functions can be formulated using mean and standard deviation values of the resulting probability distribution and general cost functions are nonlinear. To solve for the nonlinear cost function, we intentionally insert a constraint on the variance. Several candidate paths will be achieved by varying the bound value on the constraint. With fixed bound value, the Lagrangian relaxation method is applied to find the feasible solution to the constrained shortest path problem. During Lagrangian relaxation, a feasible solution close to the optimal is achieved through subgradient optimization. Among the candidate paths obtained, the best solution becomes the ultimate solution of our algorithm for the original cost function under parameter variation. The algorithm presented in this work can handle any graph structures, arbitrary edge weight distributions and general cost functions.

**Keywords** Shortest Path, Dynamic Program, Convolution, normal probability distributions.

## Introduction:

The shortest pzath problem is one of the most fundamental problems in graph theory. The objective of finding the path to minimize the cost function in the classical shortest path problem has been studied intensively. Various algorithms have been established for different implementations. A single-source, multiple-target algorithm has been developed by Dijkstra [6]. The Bellman- Ford algorithm is slower than Dijkstra's algorithm, but it is applicable for graphs with negative edge weights [7]. Another interesting algorithm is the Floyd-Warshall

algorithm which is used for multiple-source, multiple-target, or all-pair shortest path problems [8].

Recently, variation has become an important factor in analysis and has become more crucial as technology scales down. Hence, many recent researches heavily focus on improving the method and algorithm for statistical analysis. However, there are still some limitations with current works.

Research on the shortest path problem in probabilistic graphs follows work by Frank [9]. In practical situations, the costs or time is often random. This work estimates the sum of the probability distributions of the shortest path through acyclic networks weighted with random lengths. Following Frank's work, Sigal et al. [10] addressed a shortest path problem through a directed, acyclic network where arc lengths are independent random variables. This work presented an analytic derivation of path optimality indices for directed, acyclic networks.

In addition, different types of cost functions on the stochastic shortest path problem have been studied. One of the works, by Loui [11], found computationally tractable formulations of stochastic and multidimensional optimal path problems. Similar problems with maximizing the expected cost with piecewise-linear concave utility function were studied by Murthy and Sarkar [12]. Lastly, Hall [13] and Fu and Rilett [14] studied the expected shortest path on stochastic shortest path problem. To combine the previous works, X. Ji proposed three models

a- expected shortest path,

b- α-shortest path and the shortest path

c- developed a hybrid intelligent algorithm combined with genetic algorithm to solve proposed models [15].

However, these early researches on the statistical shortest path problem were designed for directed acyclic graphs (DAG) with specific edge weight distributions, such as the Gaussian distribution. Moreover, the algorithms cannot handle the general cost function, which can be nonlinear.

To overcome the limitations of previous research, Deng and Wong found an exact algorithm [16] to find the optimal solution for the cost function $\mu p + \varphi(\sigma 2p)$ on the statistical shortest path problem. Unlike former works, this

algorithm handles general graphs, arbitrary edge- weight distributions and general cost functions. To minimize the uncertainty in the final result of the **2** statistical problem, Deng and Wong added the variance constraint $\sigma 2 \leq B$ to the problem.

One of the early works that brought up the constrained shortest path problem was completed by Aneja and Nair [17]. The algorithm presented in [17] is similar to the Lagrangian multiplier technique but it requires, on average, a number of iterations which are polynomially bounded. Cai et al. [18] also studied a different approach to the constrained shortest path problem that is very similar to ours. The problem is formulated on a graph with two weight

values time and cost on a directed graph with constraints. Three variants of the problem are examined: arbitrary waiting times, zero waiting times, and vertex-dependent upper bounds on the waiting times at each vertex.

The stochastic shortest path algorithm introduced in this thesis can handle any general graph, arbitrary probability distribution, general cost function and integers as well as other non-integer values. Finally, the computation runtime is very efficient compared to previous works.

## 2. STOCHASTIC GRAPH AND OBJECTIVE

To form a graph model with variation, statistical graph structure is constructed. A pair of representative values on a probability distribution are used in objective function, and hence, assigned for edge weights. The objective function of the shortest path problem with random variables is typically nonlinear. Our approach for the nonlinear cost function is to formulate the problem into a series of constrained optimization problems with various bound values.

### 2.1 Statistical Graph

If the random variable on each edge forms a Gaussian distribution, then the overall distribution along the path, which results from adding the Gaussian distributions of each individual edge on the path, will also be Gaussian. To cover the majority of the path length distributions, the objective function may

$+¿ 3\sigma_p$ $+¿ 3\sigma_p$

be to minimize $\mu_p ¿$ or to maximize $\mu_p ¿$ . If the goal is to minimize the worst case of statistical path length from the given source node to the target

$min\ (\mu_p + 3\sigma_p)$

node, then the objective will be minimizing the cost function as $P$

for path $p$. For other general edge-weight probability distributions, the objective function can be set as Equation (1) due to Chebyshev's inequality from Equation (2).

$$\begin{array}{lll} +¿\, 3\,\sigma_p & & +¿\, 3\,\sigma_p \\ \acute{} & & \acute{} \\ \mu_p\ ¿ & & \mu_p\ ¿ \\ ¿ & & ¿ \\ ¿ & \text{or} & ¿ \qquad (1) \\ min¿ & & max\ ¿ \\ p & & p \end{array}$$

$$p(|X - \mu_p| \geq k\,\sigma_p) \leq \frac{1}{k^2} \quad \text{for random variable X and real number k > 0} \qquad (2)$$

From the above relation, any case of the given distribution can be considered by fixing the value of *k*. To solve for the stochastic shortest path problem, we only need mean $\mu$ and standard deviation $\sigma$ values from the probability distribution. This allows us to simplify the statistical graph with two real numbers on each edge weight. Mean values of the cost function probability distribution are linear functions of mean values of each edge weight distribution. However, from the Equation (3), the function to compute the overall standard deviation from multiple probability distributions is nonlinear. Hence, the conventional shortest path algorithm cannot be directly applied. Instead, overall variance can be computed linearly as shown in Equation (4). Thus, instead of considering the entire random variation, we can simply represent the statistical distribution with mean and variance values for each edge weight on the statistical graph structure as shown in **Figure 1**.

$$\sigma_p = \sqrt{\sum_{i \in p} \sigma_i^2} \quad \text{for path p} \qquad (3)$$

$$\sigma^2_p = \sum_{i \in p} \sigma_i^2 \quad \text{for path p} \qquad (4)$$

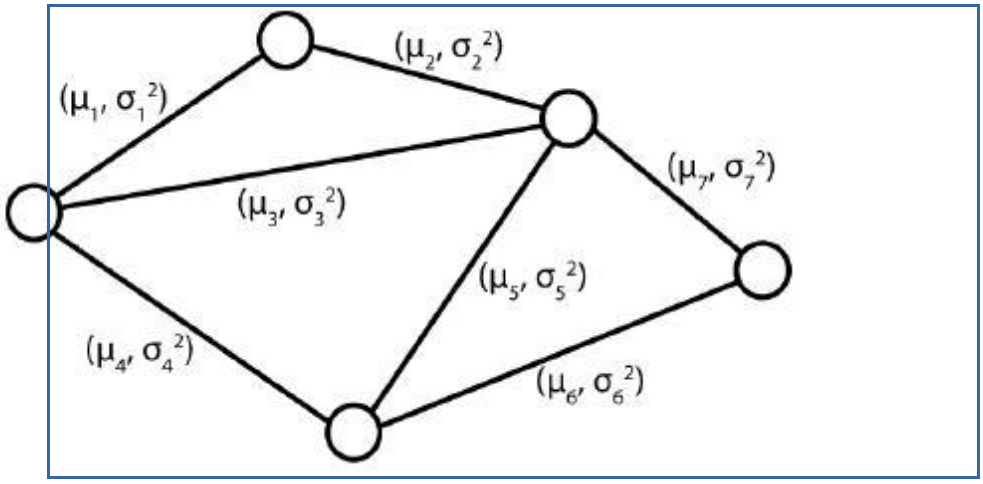


**Figure 1** Simplified statistical graph with pair $(\mu, \sigma^2)$ for each weight.

Consider the case for minimizing the worst case of any Gaussian distribution, $P$. If there is an ideal path that has a minimum mean and minimum variance, then this path will always be the optimal solution for the objective function. However, the probability that both the minimum mean and minimum variance lie on the same path is very low. Therefore it is necessary to compare the value of the cost function between paths. To achieve the goal, it will be extremely time-consuming to compare all the paths as the graph grows larger. It is sufficient to sample several path candidates and compare the goal. Typically, the candidate paths will have a small mean value but a relatively large variance, a small variance but a relatively large mean value, or fairly small values for both mean and variance.

$$min\ (\mu_p + 3\sigma_p)$$

Similarly, for maximizing the reward function of the worst case of the variation, the possibility for the existence of an ideal path with both maximum mean and maximum variance is extremely low. There are some cases for which the number of paths from the source node to target node is unbalanced. For these special structured graphs, the ideal path is likely to exist; however, the greater part would not correspond to this rare case. Path candidates for the longest path will have either a large mean value but comparatively small variance, a large variance but small mean value, or both reasonably large mean value and variance. If we negate all mean values and variances, then the objective function becomes a minimizing problem. Then, we can sample paths with properties illustrated above by applying the same method to solve for minimizing the cost function.

## 3. Problem definition and modelling

Consider a network as shown in Figure 2 consisting of a finite set of nodes and arcs of the directed acyclic network. We assume that the admissible paths are always continuous and the length of each arc is normal random variable with parameters μ and $\sigma^2$. We want to find the shortest path from the source node 1 to the sink node N using the backward dynamic programming approach.

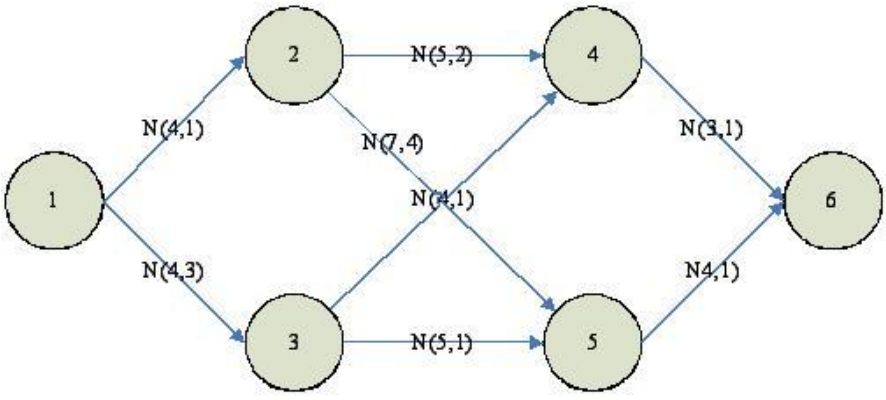


**Figure 2**. Acyclic network with normal distributed arcs

The optimal value function can be defined by

$S_i$= the distribution of the shortest path from node i to node N.

$S_i$

Then the recurrence relation can be stated as

For i = N-1,…1 (5)

$S_i \quad \min_{ji} d_{ij} \quad S_j$

And the boundary condition is

$S_N \quad 0.$

In this paper, we use convolution to find distribution of sum of two normal distributions in each stage in each stage. The reason is the inefficiency of the methods such as maximum likelihood estimation and moment generating function in long computational efforts and inaccurate solution results.

And for comparison in each stage we find the probability that a random variable with first distribution become smaller than another random variable with second distribution. In order to show the operation in each stage, represent the convolution and comparison between two normal density functions in subsequent section.

**Definition 1:** Let $X$ and $Y$ be two continuous random variables with density functions $f(x)$ and $g(y)$, respectively. Assume that both $f(x)$ and $g(y)$ are defined for all real numbers. Then the convolution $f * g$ of $f$ and $g$ is the function given by

$$(f*g)(z)=\int_{-\infty}^{+\infty} f(x)\, g(z-x)\, dx$$

¿ $\int_{-\infty}^{+\infty} f(z-y)\, g(y)\, dy$ (6)

**Theorem 1:** Let X and Y be two independent random variables with density functions fX(x) and fY(y) defined for all x. Then the sum Z = X +Y is a random variable with density function fZ(z), where fZ is the convolution of fX and fY.

$$f_z(z)=\int_{-\infty}^{+\infty} f_{X,Y}(x, z-x)\, dx$$

¿ $\int f_{X,Y}(z-y, y)dy \quad -\infty$

**Proof:** as we knew the joint density function of independent variables is equal to the products of their density functions therefore to find density function of Z = X + Y we apply cumulative distribution function technique.

$X=x$

$X+Y \leq z|$¿

¿

$P$ ¿

$$P(Z \le z) = P(X + Y \le z) = \int_{-\infty}^{+\infty} P(x + y \le z) f_X(x)\, dx = \int_{-\infty}^{+\infty} F_Y(z - x) f_X(x)\, dx$$

Now, we set partial derivative to obtain the summation density function

$$f_Z(z) = \frac{d F_Z(z)}{dz} = \frac{d}{dz}\left[\int_{-\infty}^{+\infty} F_Y(z-x) f_X(x)\, dx\right] = \int_{-\infty}^{+\infty} \frac{d F_Y(z-x)}{dz} f_X(x)\, dx = \int_{-\infty}^{+\infty} f_Y(z-x) f_X(x)\, dx \quad (8)$$

### 3.1. Sum of two independent normal random variables

Suppose that we have two random variables X and Y with a normal density function with parameters μ and $\sigma^2$. We represent the density function of Z = X + Y as follows

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_x} e^{\frac{-(x-\mu_x)^2}{2\sigma_x^2}},$$

$$f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma_y} e^{\frac{-(x-\mu_y)^2}{2\sigma_y^2}}.$$

$$f_Y(z-x) f_X(x)\, dx = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_y} e^{\frac{-(z-x-\mu_y)^2}{2\sigma_y^2}} \frac{1}{\sqrt{2\pi}\sigma_x} e^{\frac{-(x-\mu_x)^2}{2\sigma_x^2}} dx$$

+

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sqrt{\sigma_x^2+\sigma_y^2}} \exp\left[\frac{-(z-(\mu_x+\mu_y))^2}{2(\sigma_x^2+\sigma_y^2)}\right] \frac{1}{\sqrt{2\pi}\frac{\sigma_x\sigma_y}{\sqrt{\sigma_x^2+\sigma_y^2}}} \exp\left[\frac{-\left(x-\frac{\sigma_x^2(z-\mu_y)+\sigma_y^2\mu_x}{\sigma_x^2+\sigma_y^2}\right)^2}{2\frac{\sigma_x\sigma_y}{\sqrt{\sigma_x^2+\sigma_y^2}}}\right] dx$$

$$\frac{1}{\sqrt{2\pi}\sqrt{\sigma_x^2+\sigma_y^2}} \exp\left[\frac{-(z-(\mu_x+\mu_y))^2}{2(\sigma_x^2+\sigma_y^2)}\right] \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\frac{\sigma_x\sigma_y}{\sqrt{\sigma_x^2+\sigma_y^2}}} \exp\left[\frac{-\left(x-\frac{\sigma_x^2(z-\mu_y)+\sigma_y^2\mu_x}{\sigma_x^2+\sigma_y^2}\right)^2}{2\frac{\sigma_x\sigma_y}{\sqrt{\sigma_x^2+\sigma_y^2}}}\right] dx$$

$$f_Z(z)= \frac{1}{\sqrt{2\pi}\sqrt{\sigma_x^2+\sigma_y^2}} \exp\left[\frac{-(z-(\mu_x+\mu_y))^2}{2(\sigma_x^2+\sigma_y^2)}\right]$$

**Result 1:**if *X1*, *X2*,…, *Xn* are independent normal random variables with ($\mu 1,\sigma 1^2$), ($\mu 2,\sigma 2^2$),…, ($\mu n,\sigma n^2$) , then $Y = \sum_{i=1}^{n} X_i$ follows normal distribution with ( $N \sum_{i=1}^{n} \mu_i , \sum_{i=1}^{n} \sigma_i^2$ ¿ $N$ ).

### 3.2. Finding minimum density function

Now we illustrate the method that we use to find minimum between two normal random variables. In order to find the minimum random variable we compute the probability that the first random variable X1 with normal density function with ($\mu 1,\sigma 1^2$) became smaller than the second random variable X2 with normal density function with ($\mu 2,\sigma 2^2$) with considering result 1 we have

$$P(X_1 < X_2) = P(X_1 - X_2 < 0) = P\left(Z < \frac{0-(\mu_1-\mu_2)}{\sqrt{\sigma_1+\sigma_2}}\right) = \varphi\left(\frac{\mu_2-\mu_1}{\sqrt{\sigma_1+\sigma_2}}\right) \quad (10)$$

#### 3.2.1. Sum of two independent gamma random variables with same rate parameter

Suppose that we have two random variables X and Y with a gamma density function with parameter $\lambda>0$ and $\alpha>0$. We represent the density function of Z = X + Y as follows

$$f_X(x) = \frac{\lambda^{\alpha_1}}{\Gamma(\alpha_1)} x^{\alpha_1-1} e^{-\lambda x}, \quad x \geq 0,$$

$$f_Y(x) = \frac{\lambda^{\alpha_2}}{\Gamma(\alpha_2)} x^{\alpha_2-1} e^{-\lambda x}, \quad y \geq 0,$$

$$f_Z(z) = \int_{-\infty}^{+\infty} f_Y(z-x) f_X(x)\, dx = \int_0^{z} \frac{\lambda^{\alpha_2}}{\Gamma(\alpha_2)} (z-x)^{\alpha_2-1} e^{-\lambda(z-x)} \frac{\lambda^{\alpha_1}}{\Gamma(\alpha_1)} x^{\alpha_1-1} e^{-\lambda x}\, dx$$

$$= \frac{\lambda^{\alpha_1+\alpha_2}}{\Gamma(\alpha_1)\Gamma(\alpha_2)} e^{-\lambda z} \int_0^{z} (z-x)^{\alpha_2-1} x^{\alpha_1-1}\, dx$$

$$\frac{}{\Gamma(\alpha_1+\alpha} \alpha_2^{-1}$$
¿ 2) $z\,e$

As a result, if X1, X2,…, Xn are independent gamma random variables with (α1,λ), (α2,λ),…,

| | | | |
|---|---|---|---|
| | $N$ | | $n$ |
| (αn,λ) | | follows gamma distribution with λ | |
| then | $Y=\sum X_i$ | and | $\alpha=\sum \alpha_i$ . |
| | $i=1$ | | $i=1$ |

$\lambda\alpha_1+\ldots+\alpha_n$
$f_Y(y)=\frac{}{\Gamma(\alpha1+\ldots+\alpha n)} x^{\alpha1+\ldots+\alpha n+1} e^{-\lambda x}$

Now we illustrate the method that we use to find minimum between two gamma random variables see [25].

$\infty$ $\infty$ $\infty$ $\alpha_2$ $\alpha_1$
$p(X_1<X_2)=\int p(X_1<X_2|X_1=x_1$ $\frac{\lambda_2}{x}{}^{x2}2^{\alpha-1}e^{-\lambda}2\frac{\lambda_1}{}{}^{x1}x1^{\alpha}1^{-1}e^{-\lambda}1\,{}^{x}1_{d}$
$).f_{x1}$ $(x_1)\,d\,x_1=\iint \Gamma$ $\Gamma(\alpha_{1x2\,d\,x1}$
0 0 $x_1$ $(\alpha_2)$ )

## 4. Sampling Method Using Constraints

To sample the path candidates for the objective function, we will formulate the problem as a series of constrained shortest path problems shown in Equation (11). This problem finds a shortest path with minimum overall mean with respect to the variance constraint. As shown in Equation (12), varying the bound value *B* from the upper limit of variance to a lower limit

creates a number of constrained optimization problems. Solutions for each constrained problem may result in different paths.

¿
$min\ \mu_{p\ \sigma}$
$p \in P$

*subject*
¿

$p^{¿}=\{¿\ p^2 \le B\ ¿\}$
where $^{\mu}p^{=}\sum{}^{\mu}i$ ,
$i \in p$

$$min\ \sigma^2_p < B \le max\ \sigma^2_p + \mathcal{E}\ for\ small\ \mathcal{E} \quad (12)$$

Initially, the bound value will be set as the upper limit of variance; therefore, the constraint is very loose. This problem considers all the paths to obtain the shortest path regardless of the value of variance as long as it meets the constraint. As the bound value gets smaller, i.e., tighter constraint, the feasible path will have a smaller variance; however, the mean value is likely to increase due to the narrower feasible set. All these path solutions target either minimizing the mean, variance or both; hence, these are eligible as the path candidates for the original objective.

Depending on the value of the bound, some constrained problems might have the same solution path. To eliminate the redundancy, we set the bound value according to the variance of the current solution, for example assume

$\sigma 2B\ 1$ is the variance of the solution path to the constrained problem with bound value B1. Until the bound value reaches the variance of the current solution, B = $\sigma 2B\ 1$ , this path stays in the feasible set and it is the best solution. We set the next bound value to be slightly smaller than the variance of the current solution, for small B = $\sigma 2B\ 1-\mathcal{E}$ , in order to move this solution out of the feasible set.
Then, a different solution path will be achieved for the newly formulated constrained shortest path problem.

### 4.1 Lagrangian Relaxation

A well known approach for the constrained optimization problem is the Lagrangian relaxation method [19]. This method can also be applied for the shortest path problem with an additional side constraint. The basic concept of Lagrangian relaxation is to combine the constraint into the objective function by relaxing the constraint. Then, as shown in Equation (13), we can easily approach the constrained shortest path problem indirectly with the modified cost
function without any constraints, $\mu p$ + $\lambda\ (\sigma 2p-B)$. Lagrangian multiplier and it is also known as the Lagrangian multiplier problem. The Lagrangian multiplier problem is a

dual problem to the primal problem, which is the constrained optimization problem [20]. One common property of the relationship between the primal and the dual problem is weak duality. Weak duality means the optimal objective function value $g^*$ of the Lagrangian multiplier problem in Equation (8) is always a lower bound on the optimal objective function value of the primal problem $p^*$ from Equation (5),
i.e. . $g¿ \leq p¿$ With respect to the weak duality property, the solution for the

Lagrangian multiplier problem will give a lower bound value for the primal problem. In Equation (9), this weak duality relationship is manifested.

$$L(\lambda) = \min_{\lambda \geq 0} \left\{ \max_{p \in P} \left( \mu_p + \lambda \sigma_p^2 - B \right) \right\} \quad \text{for path } P$$

$$g^* = \max_{\lambda \geq 0} \left\{ \min_{p \in P} \left( \mu_p + \lambda \sigma_p^2 - B \right) \right\}$$

subject to

$$\max_{\lambda \geq 0} \min_{p \in P} \{\mu_p + \lambda (\sigma_p^2 - B)\} = g^* \leq p^* = \{ p^2 \leq B \} \text{ for path } P$$

With fixed constant $B$, the modified cost function, $\mu_p + \lambda (\sigma_p^2 - B)$, of the Lagrangian multiplier problem can be used as weights on the graph as drawn on Figure 3. The Graph in Figure 7 is converted from the example shown in Figure 4(a). Now, the graph has only single weight value and it has the conventional graph structure. On this graph, we can first apply the shortest path algorithm to solve the Lagrangian multiplier problem after fixing the Lagrangian multiplier value $\lambda$. For general graphs with positive weights, Dijkstra's shortest path algorithm can be applied. Dijkstra's algorithm is a graph search algorithm that solves the single-source shortest path problem for a graph with nonnegative edge path costs [6]. The algorithm finds the path with lowest cost between a given source vertex and every other vertex. In the case with negative weights, the Bellman-Ford algorithm can be used or if the graph is directed and acyclic, the topological sort algorithm can be used for more efficient runtime.

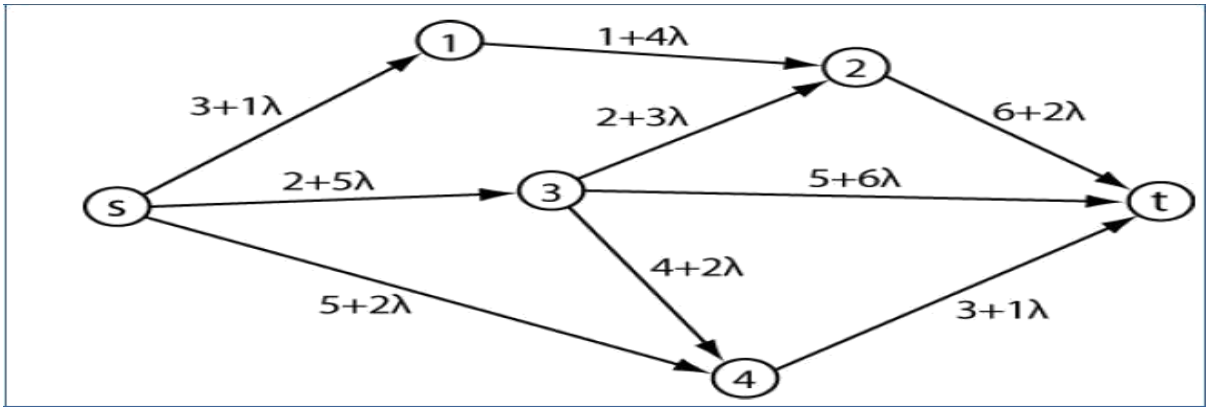


**Figure 3** Graph with modified cost function as weight.

The resulting solution of the modified cost function from the graph in Figure 3 will form a piecewise linear concave function over the Lagrangian multiplier $\lambda$. All the points on this minimum envelop of the Lagrangian function are lower bounds for the optimal solution of the original problem. Among the points, the supremum of the Lagrangian multiplier problem or dual problem, least upper bound $g^*$ which is equal to the peak value of the concave function, is the closest to the optimal and will be the best lower bound for the primal problem. The relationship between the primal and the dual problem is shown in Figure 8.

The supremum point of the dual problem can be obtained by maximizing the Lagrangian multiplier function.

### 4.2 Subgradient Optimization

The Lagrangian multiplier problem, or the dual problem, is always a concave function but not necessarily differentiable. To account for this situation, the subgradient optimization technique has been implemented to maximize the Lagrangian multiplier problem, which is nondifferentiable. Subgradient optimization is a generalization of the steepest descent method, i.e. the gradient method. The idea of this optimization technique is to move to the direction *d* where directional derivative of function *f* is positive, , with small enough step length. Equation (10) is derived from Equation (7) and since $\lambda$ is the variable, the gradient of equation (10) is shown in Equation (11). To the positive gradient direction, we move the lambda value to reach to the best value. The equation to update lambda during iteration is stated in Equation (12). From this equation, step size $\theta$ should be carefully assigned in order to guarantee the convergence.
Practitioners of the Lagrangian relaxation method often use the following heuristic for selecting the step length stated in Equation (13).

$$(\lambda)=min\ \mu + \lambda(\sigma^2-B)\ p \in P$$
$$\{p\ P \quad | \quad \}$$
$$\mathrm{L}\ p$$
$$\frac{d\,L\,(\lambda)}{d\,(\lambda)} = \sigma^2_p - B$$

$\lambda_{k+1} \leftarrow \lambda_k + \theta_k (\sigma^2_p - B)$
$\theta_k = \frac{vk\,[UB - L(\lambda k)]}{\|\sigma^2_p - B\|}$, *UB*=*upper bound* , $0<v_k \leq 2$

In the step size equation, the scalar *vk* is any number between 0 and 2. Throughout the iteration, it will be reduced by a factor of 2 whenever the best Lagrangian objective function value found so far has failed to improve in a specified number of iterations. *UB* is the upper bound of the optimal objective function. The initial upper bound can be any known feasible solution to the problem. During the subgradient iteration, the upper bound will be updated if a smaller feasible solution has been generated. A feasible path that provides the upper bound approaches the optimal path as the subgradient iteration converges. Several constrained problems formed with different bound value *B* will sample different feasible paths which are the solution candidates for the original objective function. The value of the goal will be calculated and compared among the candidate paths. Afterwards, the best result will be chosen as the solution of the Lagrangian relaxation based stochastic shortest path algorithm.

### 4.3 Algorithm

The stochastic shortest path algorithm based on the Lagrangian relaxation method described in previous sections is provided below. It is the global view of the algorithm, and the existing algorithm was implemented on shortest path. For the subgradient optimization, he initial

value of v is chosen as 0.8 after several experiments on convergence. We defined the convergence criterion to be when the error between previous and current λ values is small enough within 10 000 iterations.

Begin
Input given graph with ( $\mu_i, \sigma_i^2$ ) as weight;
Calculate min $\sigma^2 p$ and max $\sigma^2_p$ ;
Let initial UB = $\mu_p$ of the any known path;
Set initial B = max $\sigma^2_p + \varepsilon$ ;
Set L(0) ← min $\mu_p$ ;

Set v = 0.8, θ and λ; Repeat until B <min $\sigma 2p$ ,
Topological sort shortest path for min{ $\mu p + \lambda( \sigma 2p - B)$}

$\sigma_{cur}^2 \leftarrow \sigma^2_p$ of shortest path for min{ $\mu_p + \lambda(\sigma^2_p - B)$}

$\mu cur \leftarrow \mu_p$ of shortest path for min{ $\mu_p + \lambda( \sigma^2_p - B)$}

If converges,

Set next B = $\sigma_{cur}^2 - \varepsilon$ ;

Reset v = 0.8, θ and λ using

Calculate $\mu$ ¿ $+$¿ $k\sigma$ of the sampled paths with different B; $+$¿ $k\sigma$

Obtain min $\mu$ ¿ or max $\mu$ ¿ ;

Equation (12)-(13); Else,
Update v ← v/2;
Set θ and λ using Equation (12)-(13) with updated v; End if
End

## 3. Numerical example

Consider the network depicted in Figure 1. We want to obtain the shortest path from node 1 to node 6 where arcs have normal distribution.
Boundary condition is

S6 = 0.

Using the recurrence relation (1) we have

$S_5 = N(4,1)$ , $S_4 = N(3,1)$

For each arc doesn't exist in network we replace infinity for dij, so operations of S3 can be stated as

$$S3 = min\,[N(4,1)+S_5,\ N(5,1)+S_4] = min[N(4,1)+N(3,1),\ N(5,1)+N(4,1)]$$

Using result 1 we have

$$S3 = min\,[N(7,2),\ N(9,2)]$$

We find the minimum value between two normal random variables by using formula (4) as follows

$$P(X_1 < X_2) = P(X_1 - X_2 < 0) = P_{(Z < \frac{\mu - \mu}{\sqrt{\sigma_1^2 + \sigma_2^2}})} = P\left(Z < \frac{9-7}{\sqrt{2+2}}\right) = \varphi\left(\frac{2}{\sqrt{4}}\right) = \varphi(1) = 0.8413$$

So the first density function is minimal,

$S_3 = N(7,2)$

We illustrate the operation of node 2 as follows

$$S2 = min\,[N(5,2)+S_4,\ N(3,1)] = min\,[N(5,2)+N(7,4),\ N(3,1)+N(4,1)] = min[N(11,5),\ N(8,3)]$$

To find minimum density, we make use of the following probability,

$$P(X_1 < X_2) = \varphi\left(\frac{11-8}{\sqrt{3+8}}\right) = \varphi\left(\frac{3}{\sqrt{8}}\right) = \varphi\left(\frac{3}{4\sqrt{2}}\right) = 0.8556$$

So with probability 0.8556 we choose first density function as minimum density function.

$S_2 = N(8,3)$

Now we do operations for S1 to find the shortest path in network

$$P\begin{pmatrix} N(4,1)+S_2 \\ S1 = min\,[N(4,3)+S_3] \end{pmatrix} \quad \begin{matrix} N(4,1)+N(5,2)+N(3,1) \\ = min[N(4,3)+N(4,1)+N(3,1)] \end{matrix} \quad \begin{matrix} N(12,4) \\ = min[N(11,5)] \end{matrix}$$

$$X_1 < X_2) = \varphi\left(\frac{11-12}{\sqrt{9}}\right) = \varphi\left(\frac{-1}{3}\right) = 0.3694 \quad S_1 = N(11,5)$$

## CONCLUSION

As the impact of process variation increases, we cannot rely on the results from conventional algorithms that assume variations are negligible. The solution of the traditional algorithms might be very different from the statistical problem considering the variations that form a probability distribution for edge weights on the graph. There were several researches on the statistical shortest path; however, previous algorithms were not very efficient in runtime and the majority of them had limitations on practical applications. In this work, we have introduced an efficient way to obtain a reasonable solution based on Lagrangian relaxation. We intentionally insert a constraint and formulate a series of constrained problems by varying the variance bound value and sample candidate solutions for each formulated problem. At the end, we compare the objective value of the sampled candidates and attain the best solution.

The method we proposed in this work can be used in various applications in nanometer design that potentially have high parameter variations that are significant. Common applications in nanometer designs include timing analysis, maze routing, and buffer insertion. For timing analysis, precise timing information is necessary for circuit optimization to meet the yield or avoid over design. Maze routing finds the shortest path in the grid routing problem. The parameter variations cause the edge weights to be a probability distribution, and the cost functions are mostly related to the variations. Buffer insertion is a commonly used interconnection optimization technique. The possible buffer inserting location can be structured as nodes and the wire interconnection can be edges on the graph. Our algorithm can be implemented on the above listed nanometer circuit design applications to achieve efficient runtime.

**REFERENCES**


1. S. Borkar, T. Karnik, S. Narendra, J. Tschanz, A. Keshavarzi, and V. De, “Parameter variations and impact on circuits and microarchitecture,” in

2. Proceedings of the 40th Design Automation Conference, 2003, pp. 338-342.

3. Incentia Design Systems, Inc., “Advanced On-chip-variation Timing Analysis for Nanometer Designs,” Tech. Rep., Jun. 2007.

4. T. Karni, S. Borkar, and V. De, “Sub-90nm technologies—Challenges and opportunities for CAD,” in Proceedings of the International Conference on Computer Aided Design, 2002, pp. 203-206.

5. D. Pham, S. Asano, M. Bolliger, M. N. Day, H. P. Hofstee, C. Johns, J. Kahle, A. Kameyama, J. Keaty, Y. Masubuchi, M. Riley, D. Shippy, D. Stasiak, M. Suzuoki, M. Wang, J. Warnock, S. Weitzel, D. Wendel, T. Yamazaki, and K

6. Yazawa, “The design and implementation of a first-generation CELL processor,” in Proceedings of the International Solid-State Circuits Conference, 2005, pp. 184-186.

7. N. H. Chang, V. Kanevsky, O. S. Nakagawa, and S. Y. Oh, "Method and system for determining statistically based worst-case on-chip interconnect delay and crosstalk," U.S. Patent 6,018,623, Jan. 25, 2000.

8. E. W. Dijkstra, "A note on two problems in connection with graphs," Numerische Mathematik , vol. 1, pp. 269-271, 1959.

9. R. E. Bellman, "On a routing problem," Quarterly of Applied Mathematics, vol. 16, no. 1, pp. 87-90, 1958.

10. T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, Introduction to Algorithms. 2nd ed. Cambridge, MA: The MIT Press, McGraw-Hill, 2001.

11. H. Frank, "Shortest paths in probabilistic graphs," Operations Research, vol. 17, no. 4, pp. 583-599, Jul.-Aug. 1969.

12. C. E. Sigal, A. A. B. Pritsker, and J. J. Solberg, "The stochastic shortest

13. route problem," Operations Research Society of America, vol. 28, no. 5, pp. 1122-1129, Sep.-Oct. 1980.

14. R. P. Loui, "Optimal paths in graphs with stochastic or multidimensional weights," Communications of the ACM, vol. 26, no. 9, pp. 670-676, Sep. 1983.

15. Murthy and S. Sarkar, "Stochastic shortest path problems with piecewise-linear concave utility functions," Management Science, vol. 44, no. 11, pp. S125-S136, Nov. 1998.

16. R. Hall, "The fastest path through a network with random time-dependent travel time," Transportation Science, vol. 20, no. 3, pp. 182-188, Aug. 1986.

17. L. Fu and L. R. Rilett, "Expected shortest paths in dynamic and stochastic traffic networks," Transportation Research, vol. 32, no. 7, pp. 499-516, Sep. 1998.

18. [15] X. Ji, "Models and algorithm for stochastic shortest path problem,"
19. Applied Mathematics and Computation, vol. 170, no. 1, pp. 503-514, Nov. 2005.

20. L. Deng and M. D. F. Wong, "An exact algorithm for the statistical shortest path problem," in Proceedings of the Asia and South Pacific Design Automation Conference (ASP-DAC), 2006, pp. 965-970.

21. Y. P. Aneja and K. P. K. Nair, "The constrained shortest path problem,"

22.Naval Research Logistics Quarterly, vol. 25, no. 3, pp. 549-555, 1978.

23.X. Cai, T. Kloks, and C. K. Wong, "Time-varying shortest path problems with constraints," Networks, vol. 29, no. 3, pp. 141-150, Dec. 1998.

24.R. K. Ahuja, T. L. Magnanti, and J. B. Orlin, Network Flows, Theory, Algorithms, and Applications. Upper Saddle River, NJ: Prentice Hall, 1993.

25. D. P. Bertsekas, Nonlinear Programming. Belmont, MA: Athena Scientific, 1999.