Timing Failure Analysis of VLSI Circuits by Extreme Value Distribution

Meng-Chiou Wu and *Rung-Bin Lin*

Department of Computer Engineering and Science
Yuan-Ze University
Chung-Li, Taiwan, R.O.C.

Abstract

Statistically, the longest path in a VLSI circuit is not uniquely defined when process variation is considered. In fact, any long enough path could possibly becomes the longest one after fabrication. The probability distribution of the longest path delay should be collectively decided by the long enough paths. In this paper Gumbel distribution is proposed to approximate this probability distribution. The advantage over the PERT approach is that multiple long enough paths are considered and their underlying path delay distributions can be unknown. The results from experimenting with the ISCAS'85 benchmark circuits show its effectiveness. The probability distribution can be employed to compute the timing failure probability of a circuit with a given clock cycle period.

1. Introduction

Statistical timing analysis (STA) [1,2], taking into account the variation of fabrication process, provides designers with a probability distribution of the longest path delay. It is often used along with the block-oriented path tracing [3,4] to justify the timing performance of a circuit. The probability distribution can be employed to compute the probability of timing failure.

Two methods have been used to find a probability distribution of the longest path delay of a circuit. The first method is statistical timing simulation. Although it can generate a fairly accurate probability distribution of the longest path delay when a large number of experiments is performed, it is computationally intensive for a large VLSI circuit. Therefore, the past research has focused on developing efficient methods to reduce simulation time. In [5] the authors derive a set of symbolic delay expressions for the paths which may be potentially the longest one. Simulation is then performed to generate the longest path delay based on the set of symbolic delay expressions. This approach is very efficient when the number of delay expressions is small, but the assumption that the delays of two cell instances referencing the same library cell are treated as a single random variable is misleading. In [6] a part of circuit consisting of k dominant long paths is taken for timing simulation with correlation of node delay being taken into account. Since k is usually relatively smaller, compared to the number of paths in the whole circuit, the computational time is greatly reduced.

The second method is based on the PERT-like [2,7] approach which approximates the real probability distribution by the probability distribution of the path with the largest mean delay. This distribution is accurate when there are only a few dominant long paths in a circuit, not a very likely case for a high performance VLSI circuit. In [8] an approximated probability distribution was employed to find the probability of timing failure of a circuit. Statistical correlation among paths are considered, but it is computationally intensive.

Researches on the probability distribution of the longest path delay for stochastic networks [7,9] have long been performed. Among the most successful researches, PERT stands out as the most widely used technique. However, PERT may underestimate the probability of timing failure as the number of dominant long paths is large. Recent researches such as the one in [9] are proposed to tackle this problem for stochastic networks. In this paper we make an attempt to apply the result found in [9] to STA of VLSI circuits, one type of stochastic networks. Several approaches are proposed to characterize the number of dominant long paths in a circuit and the parameters of the probability distribution. The rest of this paper is organized as follows. Section 2 contains a preliminary about STA. Section 3 depicts the approach to STA based on extreme value distribution. The experimental results from the ISCAS'85 benchmark circuits are demonstrated in Section 4. Some concluding remarks are drawn in Section 5.

2. Preliminary

Although statistical approach has been used to address the problem of timing failure of a circuit, a problem formulation of STA is only recently defined explicitly in [6]. The problem is briefly stated as follows. Let G=(V,E) be a directed acyclic graph that represents a VLSI circuit. V is the set of vertices that represent the logic cells and E is the set of edges that

*This work is supported by National Science Council, Republic of China, under Grant NSC86-2615-E-155-001.

represent the signal nets connecting the logic cells. A path π in the graph G is defined as an alternating sequence of vertices and edges from one of the sources (primary inputs or memory devices) to one of the sinks (primary outputs or memory devices). Let V_π and E_π be respectively the sets of vertices and signal nets of the path. Then, the path delay T_π of path π is expressed as $T_\pi = \sum_{v \in V_\pi} S_v + \sum_{\eta \in E_\pi} N_\eta$, where S_v is the switching delay of a logic cell, and N_η is the propagation delay of a signal net. To simplify the discussion we combine cell (or vertex) delay and signal net (or edge) delay into node delay. Since long paths usually decide the timing performance of a circuit, only long paths that could possibly violate setup time are considered in this paper. The approach to short paths is similar.

The distribution of the largest (or smallest) value of a random variable is called the largest (or smallest) extreme value distribution. Thus, the problem of STA can be formulated as follows:

Given a set $\Pi = \{T_{\pi_i} | i = 1,2,\ldots,m\}$ *of random variables which represent all the possible path delays in a VLSI circuit, where* T_{π_i} *and* π_i *are defined above, find the probability distribution of the largest extreme of the statistics*

$$Y = \max_{\pi_i \in \Pi} (T_{\pi_i}), i = 1,2,\ldots,m \qquad (1)$$

and compute the timing failure probability $1 - \Pr(Y \le r)$ *for a given projected latest required arrival time r.*

3. Extreme Value Distribution Approach

Solving STA problem by simulation or multivariate analysis could need a large amount of memory and time for very large circuits. Since no closed-form solution exists so far except for some special cases, only approximate solutions are possible. The difficulties of obtaining an exact distribution for the statistics *Y* are due to dependency among paths and non-identical distribution of path delays. There are two special cases [10] that the distribution of *Y* can be obtained. First, if the path delays $T_{\pi_1}, T_{\pi_2}, \ldots, T_{\pi_m}$ are independent random variables, the probability of timing failure can be computed by

$$1 - \Pr(T_{\pi_1} \le r, T_{\pi_2} \le r, \ldots, T_{\pi_m} \le r) = 1 - \prod_{i=1}^{m} F_i(r) \qquad (2)$$

where $F_i(r) = \Pr(T_{\pi_i} \le r)$. Second, if $T_{\pi_1}, T_{\pi_2}, \ldots, T_{\pi_m}$ are independent and identical distributions with $F_i(r) = F(r)$, from (2) we obtain

$$1 - \Pr(T_{\pi_1} \le r, T_{\pi_2} \le r, \ldots, T_{\pi_m} \le r) = 1 - [F(r)]^m \qquad (3)$$

When *m* approaches infinity, the distribution of *Y*, i.e. $[F(r)]^m$, asymptotically converges to one of the extreme value distributions, Gumbel, Weibull or Frechet distribution *irregardless of the underlying path delay distribution*.

However, it is generally true that only long enough paths (dominant long paths) are viable for being the longest one. If just *k* dominant long paths ($k << m$) are considered, the statistics *Y* can be approximated as follows:

$$Y' = \max_{\pi_i \in U} (T_{\pi_i}), i = 1,2,\ldots,k \qquad (4)$$

where $U \subseteq \Pi$ is the set of $|U| = k$ dominant long paths. Thus, the timing failure probability can be rewritten as $1 - \Pr(Y' \le r)$. In general, the number of paths in a VLSI cicuit is finite and their path delay distributions are not independently and identically distributed. However, for a highly optimized circuit the dominant long paths should be more or less identically distributed and the number of dominant long paths should be sufficiently large so that they can be treated as independent ones [9]. Since each individual path delay distribution usually behaves *like* a normal one, the largest extreme for the statistics Y' is more like a Gumbel distribution [9-11] when *k* is sufficiently large.

The general form of the two-parameter Gumbel distribution function is

$$\Pr(Y' \le y') = G(y') = \exp(-e^{-b_k(y'-a_k)}) \qquad (5)$$

where a_k is the location parameter (mode) and b_k is the scale parameter. The mean μ and variance σ^2 of a Gumbel distribution can be written in terms of a_k and b_k as follows:

$$\mu = a_k + 0.57722 / b_k \quad \text{and} \quad \sigma^2 = \pi^2 / (6 b_k^2) \qquad (6)$$

The values of a_k and b_k can be computed if the underlying path delay distribution (or the parent distribution) satisfies some conditions [11]. Since the path delays are normally distributed, these conditions can be satisfied. So from [11], it follows that

$$a_k = \mu_p + \sigma_p \left((2\log k)^{1/2} - (1/2)(\log\log k + \log 4\pi) / (2\log k)^{1/2} \right) \qquad (7)$$

and

$$b_k = (2\log k)^{1/2} / \sigma_p, \qquad (8)$$

where *k* is an integer to represent the number of path random variables that are identically and independently distributed with mean μ_p and standard deviation σ_p. Note that in real applications the dominant long path delay distribution can be unknown as long as the required conditions [1] are satisfied. Thus, to use Gumbel distribution, *k*, μ_p, and σ_p must be found out. Four approaches, called *Max_μ*, *Worst_case*, *Max_*$\mu + 3\sigma$, and *Weighting*, are proposed to obtain *k*, μ_p and σ_p.

(a). ***Max_μ***: *k* is simply set equal to the ratio of the gate count in the whole circuit to the number of nodes (gates) on the path with the largest mean delay. The parameters μ_p and σ_p are respectively set equal to the mean and standard deviation of the path with the largest mean delay.

(b). ***Worst_case***: k is obtained in the same manner as that for *Max_μ*. However, the parameter μ_p and σ_p are respectively set equal to the mean and the standard deviation of the longest path delay obtained by worst-case path tracing [3,4].

(c). $\boldsymbol{Max\_\mu+3\sigma}$: k is obtained in the same manner as that for *Max_μ*. However, the parameters μ_p and σ_p are respectively set equal to the mean and standard deviation of path π_i such that $\underset{\pi_j,\ j=1,n}{Max}(u_j+3\sigma_j)=u_i+3\sigma_i$. Path π_j is among the first n longest paths with path delays $\mu_{\pi_j}+3\sigma_{\pi_j}$ greater than a delay threshold $\mu_{PERT}+1.28\sigma_{PERT}$, where μ_{PERT} and σ_{PERT} are respectively the mean and standard deviation of the longest path delay obtained by the PERT approach. If the number of paths of this sort is too large, the n's value is set to 20,000.

(d). ***Weighting***: $k=1+\frac{1}{n}\sum_{i=2}^{n} e^{(\mu_i-\mu_w)/\sigma_w}\times e^{(\sigma_i-\sigma_w)/\sigma_w}$, where μ_i and σ_i are respectively the mean and standard deviation of the path π_i which is one of the first n longest paths, obtained in the same manner as that for $Max\_\mu+3\sigma$, and σ_w is the standard deviation of the delay of the first longest path. This function takes μ_i and σ_i of the path π_i into consideration when deciding k's value. The larger the values of μ_i and σ_i are, the more contribution of the path delay of π_i to the longest path delay is. The parameters μ_p and σ_p are obtained in the same manner as that for *Worst_case*.

4. Experimental Results

The ISCAS'85 benchmark circuits are used to test the proposed approaches. They are first synthesized by Berkeley SIS system [12] with different optimization objectives based on *genlib* [13]. An algorithm [14] is then employed to extract k dominant long paths. The program for extracting k dominant long paths is written in C language and is run on UltraSparc II workstations. In our experiments, each node delay in the circuits is treated as an independent random variable. If two cell instances refer to the same cell in the library, their delays are treated as two different random variables. The node delay consists of the switching delay of a cell and the delay caused by the fanout loads. Since a cell i in *genlib* is only given with nominal delay μ_i and it is also demonstrated in [15] that the ratio of the best to worst falling time of an inverter at 6σ is 4.7, the formula $\frac{\mu_i+6\sigma_i}{\mu_i-6\sigma_i}=4.7$ is employed to compute the standard deviation σ_i of node delay. The delay distribution of a gate (cell) instance that references a library cell i is assumed to be normally distributed with $N(\mu_i,\sigma_i)$, but it can be any other distribution.

Table 1 shows the mean values and standard deviations of the distribution of the longest path delay obtained by various approaches. Compared to the PERT approach, the means of the longest path delay distributions approximated by Gumbel distributions are relatively closer to those obtained by simulation. Note that *Max_μ*, *Worst_case*, and $Max\_\mu+3\sigma$ have the same k's value for each circuit. The cumulative probability distributions (*c.d.f.*) for some circuits are depicted in Figure 1 to 2. The solid lines in the figures are the *c.d.f.* of the longest path distribution obtained by simulation. It is obvious that the tail of each Gumbel distribution is much better matched to the tail of the corresponding simulated distribution.

Table 1. Parameters of probability distribution of the longest path delay

Circuit name	Gumbel Distribution (delay time unit: ns)										PERT		Simulation	
	Max_μ			Worst_case		Max_μ+3σ		Weighting						
	k	μ	σ	μ	σ	μ	σ	k	μ	σ	μ	σ	μ	σ
C17	2	4.62	0.33	4.65	0.36	4.65	0.36	2	4.63	0.37	4.40	0.30	4.50	0.29
C499	26	34.46	0.36	35.32	0.57	35.33	0.57	6	35.16	0.60	33.00	0.71	35.37	0.58
C880	17	31.95	0.39	31.94	0.45	32.14	0.45	10	31.76	0.49	30.60	0.73	31.44	0.62
C432	10	44.85	0.62	44.94	0.66	44.94	0.66	6	44.65	0.75	43.20	1.03	44.39	0.93
C1908	19	48.29	0.46	48.34	0.56	48.46	0.51	10	48.34	0.63	46.60	0.88	48.43	0.62
C2670	28	42.83	0.44	43.02	0.48	43.02	0.48	16	42.80	0.53	41.00	0.88	42.23	0.76
C3540	28	62.68	0.49	62.68	0.54	62.80	0.62	18	62.48	0.59	60.60	1.00	62.47	0.71
C5315	50	59.59	0.42	60.06	0.51	60.12	0.54	28	59.80	0.55	57.50	0.91	59.23	0.69
C1355	25	38.27	0.31	38.84	0.46	38.85	0.46	25	38.85	0.46	37.00	0.62	39.42	0.49
C7552	89	53.83	0.41	54.22	0.48	54.22	0.48	53	54.01	0.51	51.40	0.97	53.22	0.67
C6288	19	175.0	0.79	174.8	0.93	175.0	0.92	21	174.8	0.92	172.1	1.51	177.0	1.25

Table 2 shows the timing success probability for each distribution by giving a latest required arrival time, set equal to $\mu+3\sigma$ of the longest path obtained by the PERT approach. The timing failure probability of each circuit is equal to $1-\Pr(Y'\le\mu+3\sigma)$, where $\Pr(Y'\le\mu+3\sigma)$ is the timing success probability. *Worst_case*, $Max\_\mu+3\sigma$ and *Weighting* give the best prediction of the timing failure probability in average. Especially, *Worst_case* is quite efficient because it only needs to find the longest path by worst-case path tracing.

Table 2. Timing success probability at $\mu+3\sigma$ of PERT

Circuit name	PERT (μ +3σ)	Simulation	Gumbel distribution			
			Max_μ	Worst_case	Max_μ +3σ	Weighting
C17	99.87%	99.63%	96.12%	94.74%	94.74%	94.37%
C499	99.87%	36.31%	95.14%	42.34%	41.32%	55.16%
C880	99.87%	98.23%	96.26%	95.09%	91.42%	96.20%
C432	99.87%	97.96%	97.12%	96.04%	96.04%	96.63%
C1908	99.87%	90.57%	95.97%	93.03%	92.33%	91.30%
C2670	99.87%	96.77%	94.89%	89.78%	89.78%	93.03%
C3540	99.87%	93.61%	94.86%	93.56%	89.70%	95.21%
C5315	99.87%	92.14%	92.23%	68.40%	64.24%	80.67%
C1355	99.87%	12.07%	95.23%	58.77%	58.18%	58.52%
C7552	99.87%	94.02%	87.70%	63.52%	63.52%	76.33%
C6288	99.87%	38.72%	95.89%	95.48%	93.72%	95.34%
Mean	99.87%	77.27%	94.67%	80.98%	79.54%	84.80%

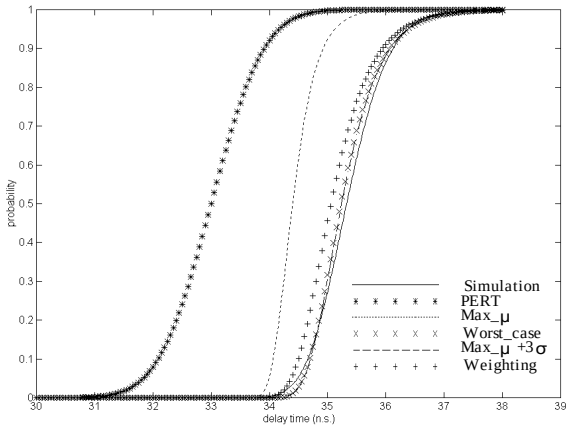

Figure 1. Cumulative distribution functions of the longest path delay of C499

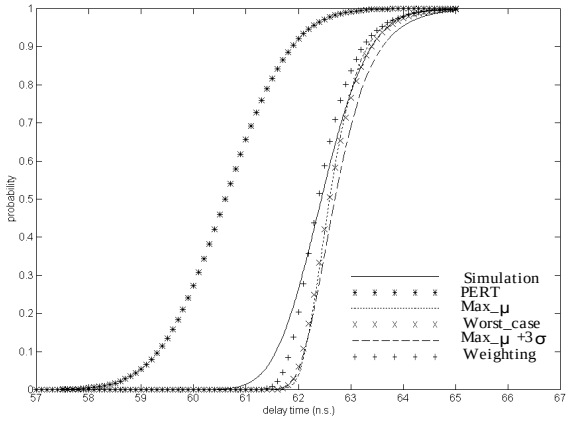

Figure 2. Cumulative distribution functions of the longest path delay of C3540

6. Concluding Remarks and Future Work

Although it is found that the average timing failure probability predicted by Gumbel distribution for a given latest required arrival time, set to $\mu + 3\sigma$ of the longest path delay obtained by the PERT approach, is in average better than that is predicted by the PERT approach. However, for some of the circuits such as C1355, C7522 and C6288, the proposed approaches do not performed well. This may be caused by the fact that path correlation is not considered in our implementation. What still left unconsidered in our current implementation is false path problem [16] and correlation of node delays. In the future, we will work to see how false paths and the correlation of node delays and path delays would influence the effectiveness of the extreme value approach.

References:

[1] T. I. Kirkpatrick and N. R. Clark, "PERT as an Aid to Logic Design," IBM J. RES. DEVELOP., March, 1966, pp. 135-141.

[2] Jacques Benkoski and Andrzei J. Strojwas, "New Techniques for Statistical Timing Simulation," the ACM/IEEE 23rd Design Automation Conference, 1986, pp. 134-137.

[3] Robert B. Hitchcock, Sr., Gordon L. Smith and David D. Cheng, "Timing Analysis of Computer Hardware," IBM J. RES. DEVELOP., Vol. 26, No. 1, January 1982, pp. 100-105.

[4] Lionel C. Bening and Thomas A. Lane, "Development in Logic Network Path Delay Analysis," the IEEE 19th Design Automation Conference, 1982, pp. 605-615.

[5] Horng-Fei Jyu, Sharad Malik, Srinivas Devadas and kurt W. Keutzer, "Statistical Timing Analysis of Combinational Logic Circuits," IEEE Trans. on VLSI, Vol. 1, No. 2, June 1993, pp. 126-137.

[6] Rung-Bin Lin and Meng-Chiou Wu, "A New Statistical Approach to Timing Analysis of VLSI Circuits," Proceedings of the 11th International Conference on VLSI Design, Chennai, India, January 1998, pp. 507-513.

[7] D. G. Malcolm, J. H. Roseboom, C. E. Clark and W. Fazar, "Application of a Technique for Research and Development Program Evaluation," Operations Research 7, 646-669 (1959).

[8] D. R. Tryon, F. M. Armstrong and M. R. Reiter, "Statistical Failure Analysis of System Timing," IBM J. RES. DEVELOP, Vol. 28, No. 4, July, 1984, pp. 340-355.

[9] Bajis Dodin and Mete Sirvanci, "Stochastic Networks and the Extreme Value Distribution," Computer Operations Research, Vol. 17, No. 4, 1990, pp. 397-409.

[10] H. A. David, "Order Statistics," John Wiley & Sons. Inc. 1981.

[11] Enrique Castillo, "Extreme Value Theory in Engineering," Academic Press, Inc. 1988.

[12] E. M. Sentovich, K. J. Singh, L. Lavagno, C. Moon, R. Murgai, A. Saldanha, H. Savoj, P. R. Stephan, R. K. Brayton, and A. Sangiovanni-Vincentelli, "SIS: A System for Sequential Circuit Synthesis," Department of Electrical Engineering and Computer Science University of California, Berkeley, CA 94720, May 1992.

[13] R. Lisanke, "Logic synthesis and optimization benchamrk: User's guide," 1989.

[14] Steve H.C. Yen, David H.C. Du, and S. Ghanta, "Efficient Algorithms for Extracting the k Most Critical Paths in Timing Analysis," 26th ACM/IEEE Design Automation Conference, 1989.

[15] M. G. Buehler, B. R. Blaes, H. R. Sayah and U. Lieneweg, "Parameter Distribution for Complex VLSI Circuits," 6th MIT Conference in Advanced Research in VLSI, 1989, pp. 159-174.

[16] J. Benkoski, E. Vanden Meesch, L. Claesen, and H. De Man. "Efficient Algorithms for Solving the False Path Problem in Timing Verification," IEEE International Conference on Computer-Aided Design, 1987, pp. 44-47.