Optimal Non-Uniform Wire-Sizing under the Elmore Delay Model

Chung-Ping Chen, Hai Zhou, and D. F. Wong
Department of Computer Sciences
University of Texas at Austin
Austin, Texas 78712

Abstract

We consider non-uniform wire-sizing for general routing trees under the Elmore delay model. Three minimization objectives are studied: 1) total weighted sink-delays; 2) total area subject to sink-delay bounds; and 3) maximum sink-delay. We first present an algorithm NWSA-wd for minimizing total weighted sink-delays based on iteratively applying the wire-sizing formula in [1]. We show that NWSA-wd always converges to an optimal wire-sizing solution. Based on NWSA-wd and the Lagrangian relaxation technique, we obtained two algorithms NWSA-db and NWSA-md which can optimally solve the other two minimization objectives. Experimental results show that our algorithms are efficient both in terms of runtime and storage. For example, NWSA-wd, with linear runtime and storage, can solve a 6201-wire-segment routing-tree problem using about 1.5-second runtime and 1.3-MB memory on an IBM RS/6000 workstation.

1 Introduction

As VLSI technology continues to scale down, interconnect delay has become the dominant factor in deep submicron designs. As a result, wire-sizing plays an important role in achieving desirable circuit performance. Almost all of the existing wire-sizing algorithms (e.g. [3, 2, 5, 6, 11]) size each wire segment uniformly, i.e., identical width at every position on the wire segment. In order to achieve non-uniform wire-sizing, these algorithms have to chop wire segments into large number of small segments. Consequently, the number of variables in the optimization problem is increased substantially and thus results in long runtime and large storage.

In this paper, we consider non-uniform wire-sizing for general routing trees. Previous works on efficient algorithms for non-uniform wire-sizing can be found in [1, 9, 12]. It was shown in [1, 9] that the wire-sizing function for a single wire (of length $\mathcal{L}$) that minimizes Elmore delay must be a decreasing function of the form ae^{-bx}, $0 \leq x \leq \mathcal{L}$, i.e. the wire-width at position x is ae^{-bx}. If lower and upper bounds of the wire-widths are given, it was shown in [1] that the optimal wire-sizing function is a truncated version of ae^{-bx}. In this paper, we show that the results in [1] can be applied to general routing trees. Three minimization objectives are studied: 1) total weighted sink-delays; 2) total area subject to sink-delay bounds; and 3) maximum sink-delay. We first present an algorithm NWSA-wd for minimizing total weighted sink-delays based on iteratively applying the wire-sizing formula in [1]. We show that NWSA-wd always converges to an optimal wire-sizing solution. Based on NWSA-wd and the Lagrangian relaxation technique [8], we obtained two algorithms NWSA-db and NWSA-md which can optimally solve the other two minimization objectives. Experimental results show that our algorithms are efficient both in terms of runtime and storage. For example, NWSA-wd, with linear runtime and storage, can solve a 6201-wire-segment routing-tree problem using about 1.5-second runtime and 1.3-MB memory on an IBM RS/6000 workstation.

The remainder of this paper is organized as follows. In Section 2, we present the notations to be used in this paper. In Section 3, we show how to compute the Elmore delay for non-uniformly sized wire segments in a routing tree. Section 4 presents the algorithm NWSA-wd for minimizing weighted delay at the sinks. Sections 5 presents the algorithm NWSA-db for minimizing area subject to delay bounds and the algorithm NWSA-md for minimizing the maximum delay at the sinks. Finally, we present some experimental results in Section 6.

2 Preliminaries

We use the following notations in this paper.

- T: A routing tree with a driver w_0 at the root (source) and a set of s sinks $\{N_1, N_2, ..., N_s\}$.
- w_i: i-th wire segment, $1 \leq i \leq n$.
- f_i: Wire-sizing function of w_i.
- d_i: Delay of w_i, $0 \leq i \leq n$.
- $\mathcal{L}_i$: Length of w_i.
- **f**: $\mathbf{f} = (f_1, f_2, \ldots, f_n)$ is a wire-sizing solution.
- ρ: Resistance of wire per unit length at unit width.
- ε: Area capacitance of wire per unit square.
- R_d: Driver resistance.
- r_i: Resistance of w_i; $r_i \approx \int_0^{\mathcal{L}_i} \rho / f_i(x)dx$, $1 \leq i \leq n$.
- c_i: Capacitance of w_i; $c_i \approx \int_0^{\mathcal{L}_i} \varepsilon f_i(x)dx$, $1 \leq i \leq n$.
- U_i, L_i: Upper bound and lower bound of the size of w_i, respectively, i.e., $L_i \leq f_i(x) \leq U_i$, $1 \leq i \leq n$.
- P_i: Driver and all wires on the path from the source to sink N_i (including N_i).
- $parent(w_i)$: Parent of w_i.
- $Child(w_i)$: *Set* of w_i's children.

- $Ans(w_i)$: Driver and all the wires on the path from the source to w_i (excluding w_i).
- $Dec(w_i)$: All the wires on the paths from w_i to the sinks (excluding w_i).
- λ_i: Sink weight associated with N_i, $1 \le i \le s$.
- μ_i: Edge weight associated with w_i, $1 \le i \le s$, where $\mu_i = \sum_{N_j \in Dec(w_i)} \lambda_j$.
- R_i: Weighted upstream resistance of w_i; $R_i = \sum_{w_j \in Ans(w_i)} \mu_j r_j$.
- C_i: Downstream capacitance of w_i; $C_i = \sum_{w_j \in Dec(w_i)} c_j + \sum_{N_j \in Dec(w_i)} \tilde{c}_j$, where $\tilde{c}_j$ is the capacitance of sink N_j, $1 \le j \le s$.
- A: Area of T; $A = \sum_{i=1}^{n} \int_0^{\mathcal{L}_i} f_i(x)dx$.

Figure 1 illustrates some of the above notations.

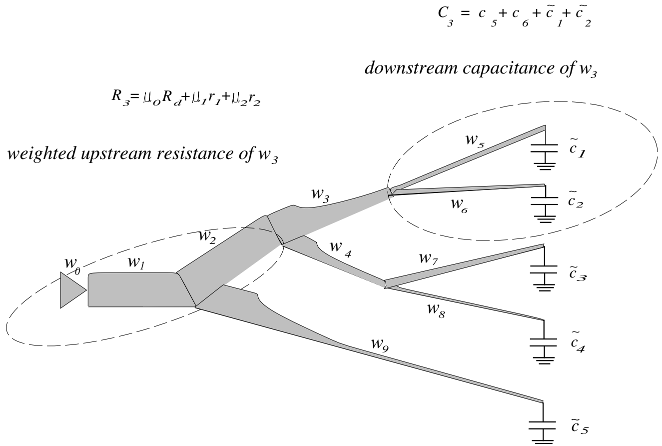

Figure 1: A non-uniformly sized routing tree T with nine wire segments and five sinks

3 Elmore Delay Model

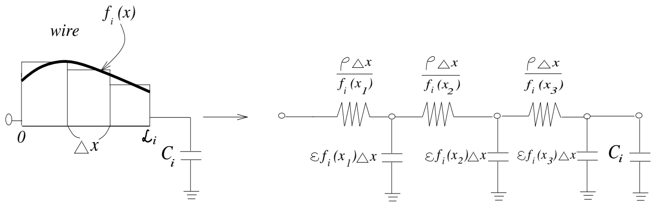

Figure 2: Elmore delay model

We use the Elmore delay model to calculate the signal delay through wire segments [4]. Suppose wire segment w_i is partitioned into m equal-length wire segments, each of length $\triangle x = \frac{\mathcal{L}_i}{m}$. Let x_j be $j\triangle x$, $1 \le j \le m$. The capacitance and resistance of wire segment j can be approximated by $\varepsilon \triangle x f_i(x_j)$ and $\rho\triangle x / f_i(x_j)$, respectively. Thus the Elmore delay through w_i can be approximated by

$$d_i = \sum_{j=1}^{m} \frac{\rho \Delta x}{f_i(x_j)} \left(\sum_{k=j}^{m} \varepsilon f(x_k) \Delta x + C_i \right)$$

It is the sum of the delay in each wire segment j, which is given by its own resistance $\rho\triangle x / f_i(x_j)$ multiplied by its downstream capacitance $\sum_{k=j}^{m} \varepsilon f_i(x_k)\triangle x + C_i$ (See Figure 2). Let $m \to \infty$, we get

$$d_i = \int_0^{\mathcal{L}_i} \frac{\rho}{f_i(x)} \left(\int_x^{\mathcal{L}_i} \varepsilon f_i(t)\, dt + C_i \right) dx \qquad (1)$$

as the Elmore delay through w_i. The driver delay d_0 is approximated by $R_d(\sum_{i=1}^{n} \varepsilon \int_0^{\mathcal{L}_i} f_i(x)dx + \sum_{j=1}^{s} \tilde{c}_j)$. Let D_i be the delay at sink N_i, we have

$$\begin{aligned} D_i &= d_0 + \sum_{w_j \in P_i \setminus \{w_0\}} d_j \\ &= R_d \left(\sum_{i=1}^{n} \varepsilon \int_0^{\mathcal{L}_i} f_i(x)dx + \sum_{j=1}^{s} \tilde{c}_j \right) + \\ &\quad \sum_{w_j \in P_i \setminus \{w_0\}} \int_0^{\mathcal{L}_j} \frac{\rho}{f_j(x)} \left(\int_x^{\mathcal{L}_j} \varepsilon f_j(t)\, dt + C_j \right) dx \qquad (2) \end{aligned}$$

In the case where there is only one wire (i.e., $n = s = 1$), the delay at the sink can be calculated as follows:

$$\begin{aligned} D_1 &= \int_0^{\mathcal{L}_1} \frac{\rho}{f_1(x)} \left(\int_x^{\mathcal{L}_1} \varepsilon f_1(t)\, dt + \tilde{c}_1 \right) dx + \\ &\quad R_d \int_0^{\mathcal{L}_1} \varepsilon f_1(x)dx + R_d \tilde{c}_1 \qquad (3) \end{aligned}$$

4 Minimizing Weighted Delay

In this section, we consider the following problem:

$$\begin{aligned} \mathcal{M}1 : \textit{Minimize} \quad & D(\mathbf{f}) = \sum_{i=1}^{s} \lambda_i D_i \\ \textit{Subject to} \quad & L_i \le f_i \le U_i,\ 1 \le i \le n \end{aligned}$$

4.1 Algorithm

We have designed an algorithm called NWSA-wd to optimally solve $\mathcal{M}1$. NWSA-wd is a greedy algorithm based on iteratively re-sizing the wire segments. Initially we set the width of each wire segment to be its minimum possible wire size, i.e. $f_i = L_i, 1 \le i \le n$. In each iteration, we traverse the routing tree T top down and examine the wire segments one at a time. Each time when we visit a wire segment w_i, we re-size it optimally while keeping the wire-sizing functions at all the other wire segments fixed, i.e. we compute f_i to minimize $D(\mathbf{f})$ while keeping f_j fixed for all $j \ne i$. We now show how to optimally re-size a wire segment.

Lemma 1 *For each i,*

$$\begin{aligned} D(\mathbf{f}) &= \mu_i \int_0^{\mathcal{L}_i} \frac{\rho}{f_i(x)} \left(\int_x^{\mathcal{L}_i} \varepsilon f_i(t)\, dt + C_i \right) dx + \\ &\quad R_i \int_0^{\mathcal{L}_i} \varepsilon f_i(x)\, dx + \xi_i(\mathbf{f}) \qquad (4) \end{aligned}$$

where $\xi_i(\mathbf{f})$ is independent of f_i.

Proof: $D(\mathbf{f})$ can be rewritten as follows:

$$\begin{aligned} D(\mathbf{f}) &= \sum_{i=1}^{s} \lambda_i D_i \\ &= \sum_{i=1}^{s} \lambda_i \sum_{w_k \in P_i} d_k \\ &= \sum_{w_k \in T} d_k \sum_{N_j \in Dec(w_k)} \lambda_j \\ &= \sum_{w_k \in T} \mu_k d_k \end{aligned}$$

Note that the terms that involve f_i come from $\sum_{w_k \in Ans(w_i)} \mu_k d_k$ and $\mu_i d_i$. Also note that, for $w_k \in Ans(w_i)$, the only term in d_k that involves f_i is $r_k c_i$. Therefore, we have

$$\begin{aligned} D(\mathbf{f}) &= \mu_i d_i + \sum_{w_k \in Ans(w_i)} \mu_k r_k c_i + \xi_i(\mathbf{f}) \\ &= \mu_i d_i + c_i R_i + \xi_i(\mathbf{f}) \end{aligned}$$

where $\xi_i(\mathbf{f})$ is independent of f_i. The theorem follows from the fact that d_i is given by (1) and $c_i = \int_0^{\mathcal{L}_i} \varepsilon f_i(x)\, dx$. □

Theorem 1 *For each w_i, the optimal re-sizing of w_i can be obtained by treating w_i as a single wire of length $\mathcal{L}_i$ with driver resistance R_i, capacitance load C_i, wire resistance per unit length at unit width $\mu_i \rho$, and capacitance of wire per unit square ε.*

Proof: Let $\xi_i(\mathbf{f})$ be defined as in Lemma 1. Since $\xi_i(\mathbf{f}), R_i$ and C_i are independent of f_i, therefore, the problem of re-sizing w_i while keeping all other wire-sizing functions fixed is equivalent to finding f_i to minimize

$$\begin{aligned} \Phi(f_i) &= D(\mathbf{f}) - \xi_i(\mathbf{f}) + R_i C_i \\ &= \int_0^{\mathcal{L}_i} \frac{\mu_i \rho}{f_i(x)} \left(\int_0^{\mathcal{L}_i} \varepsilon f_i(x) dt + C_i \right) dx + \\ & \quad R_i \int_0^{\mathcal{L}_i} \varepsilon f_i(x) dx + R_i C_i \end{aligned} \tag{5}$$

Comparing (5) with (3), we observe that $\Phi(f_i)$ is also the Elmore delay of a single wire of length $\mathcal{L}_i$ with wire-sizing function f_i, capacitance load C_i, driver resistance R_i, wire resistance per unit length at unit width $\mu_i \rho$, and area capacitance of wire per unit square ε. □

The problem of determining a wire-sizing function for a single wire to minimize Elmore delay has been solved in [1]. Theorem 1 allows us to directly apply the results in [1] to optimally re-size each wire segment. Let f_i be the wire-sizing function that optimally re-size w_i. It follows from [1] that

$$f_i(x) = \begin{cases} U_i & 0 \le x \le l_1 \\ a_i e^{-b_i(x-l_1)} & l_1 \le x \le l_1 + l_2 \\ L_i & l_1 + l_2 \le x \le l_1 + l_2 + l_3 = \mathcal{L}_i \end{cases}$$

where a_i, b_i, l_1, l_2, and l_3 can be computed in $O(1)$ time (assuming that R_i, C_i, and μ_i have been computed).

4.2 Analysis

In this section, we show that NWSA-wd always converges to an optimal wire-sizing solution. We also show that each iteration of NWSA-wd runs in linear time and uses linear storage.

Lemma 2 *During the execution of NWSA-wd, whenever we re-size w_i from f_i to f_i', we have $f_i(x) \le f_i'(x), 0 \le x \le \mathcal{L}_i$.*

Proof: We only present a sketch of the proof.

Suppose $f_i(x) = a_i e^{-b_i x}$ and $f_i' = a_i' e^{-b_i' x}, 0 \le x \le \mathcal{L}_i$. It follows from [1] and Theorem 1 that

$$a_i = \frac{\mu_i \rho}{b_i R_i}, b_i \sqrt{\frac{R_i C_i}{\mu_i \rho \varepsilon}} = e^{-\frac{b_i \mathcal{L}_i}{2}},$$

$$a_i' = \frac{\mu_i \rho}{b_i' R_i'}, \text{ and } b_i' \sqrt{\frac{R_i' C_i'}{\mu_i \rho \varepsilon}} = e^{-\frac{b_i' \mathcal{L}_i}{2}}$$

where R_i and R_i' (C_i and C_i') are the weighted up-stream resistance (down-stream capacitance) before the two consecutive re-sizing of w_i. Without of loss of generality, we may assume that all previous re-sizing operations produced larger wire-sizing functions. Therefore, we have $R_i' \le R_i$ and $C_i' \ge C_i$. Based on the above, we can prove that

$$a_i \le a_i' \text{ and } a_i e^{-b_i \mathcal{L}_i} \le a_i' e^{-b_i' \mathcal{L}_i} \tag{6}$$

Furthermore, based on (6), we can prove that

$$f_i(x) = a_i e^{-b_i x} \le a_i' e^{-b_i' x} = f_i'(x), 0 \le x \le \mathcal{L}_i$$

For the case where f_i and f_i' are truncated exponential functions, the proof can be reduced to the above special case and the details will not be presented here. □

Theorem 2 *NWSA-wd always converges to an optimal wire-sizing function.*

Proof: Let f_{ij} be the wire-sizing function of w_i after the jth iteration. It follows from Lemma 2 that $f_{i1}, f_{i2}, f_{i3}, \ldots$ is monotonically increasing. Moreover, all f_{ij}'s are upper bounded by U_i, therefore, $\{f_{ij}\}$ converges as $j \to \infty$. Let $\tilde{f}_i = \lim_{j \to \infty} f_{ij}$ and $\tilde{\mathbf{f}} = (\tilde{f}_1, \tilde{f}_2, \ldots, \tilde{f}_n)$. Clearly NWSA-wd converges to $\tilde{\mathbf{f}}$ and $\tilde{\mathbf{f}}$ is a local optimal solution. Consider the transformation $f_i = e^{g_i}, 1 \le i \le n$. Let $\mathbf{g} = (g_1, g_2, \ldots, g_n)$ and $e^{\mathbf{g}} = (e^{g_1}, e^{g_2}, \ldots, e^{g_n})$. We have $\hat{D}(\mathbf{g}) = D(\mathbf{f}) = D(e^{\mathbf{g}})$. Note that $\mathbf{f}$ is a global minimum for D if and only if $\mathbf{g}$ is a global minimum for $\hat{D}$. By expanding the terms in (2), we can show that

$$D(\mathbf{f}) = \sum_{k=1}^{m} \alpha_k h_k(\mathbf{f})$$

where $\alpha > 0$ and $h_k(\mathbf{f})$ is of the form

$$\int_0^{\mathcal{L}_i} f_i(x)\, dx, \int_0^{\mathcal{L}_j} \frac{dx}{f_j(x)}, \text{ or } \int_0^{\mathcal{L}_j} \int_x^{\mathcal{L}_i} \frac{f_i(t)}{f_j(x)}\, dt\, dx.$$

Therefore, $h_k(e^{\mathbf{g}})$ is of the form

$$\int_0^{\mathcal{L}_i} e^{g_i(x)}\, dx,\ \int_0^{\mathcal{L}_j} e^{-g_j(x)}\, dx,\ \text{or}\ \int_0^{\mathcal{L}_j}\int_x^{\mathcal{L}_i} e^{g_i(t)-g_j(x)}\, dt\, dx.$$

Since $e^{\alpha y+(1-\alpha)z} \le \alpha e^y + (1-\alpha)e^z, \forall y,z \in \mathcal{R}, 0 \le \alpha \le 1$, and integration is a linear operator, therefore

$$h_k(e^{\alpha \mathbf{g}'+(1-\alpha)\mathbf{g}''}) \le \alpha h_k(e^{\mathbf{g}'}) + (1-\alpha) h_k(e^{\mathbf{g}''}),$$

$0 \le \alpha \le 1, \forall \mathbf{g}', \mathbf{g}''$, i.e., $h_k(e^{\mathbf{g}})$ is convex in $\mathbf{g}$. Hence, $\hat{D}(\mathbf{g}) = \sum_{k=1}^m \alpha_k h_k(e^{\mathbf{g}})$ is convex in $\mathbf{g}$. Since $\tilde{\mathbf{f}}$ is a local minimum for D, therefore $\tilde{\mathbf{g}}$ is a local minimum for $\hat{D}$. Since $\hat{D}$ is convex, therefore $\tilde{\mathbf{g}}$ is a global minimum. Thus $\tilde{\mathbf{f}}$ is a global minimum for D. □

Algorithm: NWSA-wd
A1. $f_i = L_i,\ 0 \le i \le n$.
A2. Compute all μ_i's in a bottom-up traversal of T using the formula: $\mu_i = \sum_{w_j \in Child(w_i)} \lambda_j$.
A3. Compute all C_i's in a bottom-up traversal of T using the formula: $C_i = \sum_{w_j \in Child(w_i)} (C_j + c_j)$.
A4. Traverse the tree top-down:
For each w_i,
$R_i = R_{parent(w_i)} + \mu_{parent(w_i)} r_{parent(w_i)}$;
Compute f_i according to [1] and Theorem 1.
A5. Repeat **A3**–**A4** until no improvement.

Figure 3: Non-uniform wire-sizing algorithm for minimizing total weighted delays

Figure 3 shows the algorithm NWSA-wd. We now show that each iteration of NWSA-wd runs in $O(n)$ time using $O(n)$ storage. Note that we can pre-compute the set of all edge weights μ_i's in $O(n)$ time by a bottom-up traversal of T (**A2**). At the beginning of each iteration, we compute all downstream capacitances C_i's by a bottom-up traversal of T in $O(n)$ time (**A3**). Each iteration of NWSA-wd consists of a top-down traversal of T; each time we visit a wire segment w_i (**A4**), we compute f_i according to [1] as described earlier. Note that f_i can be computed in $O(1)$ time provided that we have the current values of R_i and C_i. Observe that $R_i = R_{parent(w_i)} + \mu_{parent(w_i)} r_{parent(w_i)}$ and hence can be updated in $O(1)$ time. Since C_i is independent of the wire-sizing functions of the ancestors of w_i, there is no need to update C_i during the top-down traversal of T. From the above discussion, it is clear that NWSA-wd runs in linear time and uses linear storage.

5 Minimizing Other Objectives

In this section, we present algorithms NWSA-db and NWSA-md for minimizing area subject to delay bounds and minimizing maximum delay, respectively. Both algorithms are based on the Lagrangian relaxation technique [8] while using NWSA-wd as a subroutine.

Lagrangian relaxation is a technique for solving constrained optimization problems. "Troublesome" constraints are "relaxed" and incorporated into the objective function after multiplying them by constants called Lagrange multipliers, one multiplier for each constraint. For each fixed set of Lagrange multipliers, we have a new optimization problem (free of troublesome constraints) called the Lagrangian relaxation subproblem (LRS) associated with this set of multipliers. We iteratively adjust the Lagrange multipliers and solve the corresponding LRS until the process converges. We shall show that for minimizing area subject to delay bounds and minimizing maximum delay, each LRS is equivalent to minimizing weighted delay with the Lagrange multipliers being the sink weights and hence can be solved by using NWSA-wd.

5.1 Minimizing Area

Let B_i be the delay bound at sink $N_i, 1 \le i \le s$. The problem of minimizing area subject to delay bounds can be formulated as follows.

$$\begin{array}{lll} \mathcal{M}2: \mathit{Minimize} & A & \\ \quad\quad \mathit{Subject\ to} & D_i(\mathbf{f}) \le B_i, & 1 \le i \le s \\ & L_i \le f_i \le U_i, & 1 \le i \le n \end{array}$$

Following the Lagrangian relaxation procedure, we introduce a Lagrange multiplier λ_i for each delay constraint $D_i(\mathbf{f}) \le B_i, 1 \le i \le s$. The following is the LRS associated with the Lagrange multipliers $\lambda_1, \lambda_2, \ldots, \lambda_s$.

$$\begin{array}{ll} \mathcal{M}2': \mathit{Minimize} & D'(\mathbf{f}) = A + \sum_{i=1}^s \lambda_i (D_i(\mathbf{f}) - B_i) \\ \quad\quad \mathit{Subject\ to} & L_i \le f_i \le U_i,\ 1 \le i \le n \end{array}$$

Let $\lambda = \sum_{i=1}^s \lambda_i$ and $\hat{D}_i$ be the delay at sink N_i if the driver resistance is increased to $R_d + \frac{1}{\varepsilon\lambda}$. Replacing R_d by $R_d + \frac{1}{\varepsilon\lambda}$ in (2), we get

$$\hat{D}_i = D_i + \frac{A}{\lambda} + \frac{1}{\varepsilon\lambda}\sum_{j=1}^s \tilde{c}_j$$

Therefore, the weighted delay $\hat{D}(\mathbf{f})$ is given by

$$\begin{aligned} \hat{D}(\mathbf{f}) &= \sum_{i=1}^s \lambda_i \hat{D}_i \\ &= D(\mathbf{f}) + A + \frac{1}{\varepsilon}\sum_{j=1}^s \tilde{c}_j \\ &= D'(\mathbf{f}) + \sum_{i=1}^s \lambda_i B_i + \frac{1}{\varepsilon}\sum_{j=1}^s \tilde{c}_j \end{aligned}$$

Since $\sum_{i=1}^s \lambda_i B_i + \frac{1}{\varepsilon}\sum_{j=1}^s \tilde{c}_j$ is a constant, therefore minimizing $D'(\mathbf{f})$ is equivalent to minimizing $\hat{D}(\mathbf{f})$. Thus we have the following theorem.

Theorem 3 *Solving the LRS $\mathcal{M}2'$ is equivalent to solving $\mathcal{M}1$ with driver resistance changed to $R_d + \frac{1}{\varepsilon\lambda}$.* □

Algorithm: NWSA-db
A1. $k = 0$; $f_i = L_i$, $0 \le i \le n$.
A2. $\lambda_i = 1/s$, $1 \le i \le s$.
A3. Call NWSA-wd to solve the LRS associated with $\lambda_1, \ldots, \lambda_s$.
A4. Recursively compute D_i, $1 \le i \le s$.
A5. $\lambda_i = \lambda_i + \theta_k(D_i - B_i)$, $1 \le i \le s$.
A6. $k = k + 1$.
A7. Repeat **A3**–**A6** until no improvement.

Figure 4: Non-uniform wire-sizing algorithm for minimizing total area subject to delay bounds

Figure 4 shows the algorithm NWSA-db for solving $\mathcal{M}2$. We update the Lagrange multipliers in **A5** where the sequence $\{\theta_k\}$ satisfies $\lim_{k\to\infty} \theta_k = 0$ and $\sum_{j=1}^{k} \theta_j \to \infty$ as $k \to \infty$. It can be shown that NWSA-db converges to an optimal wire-sizing solution (for $\mathcal{M}2$).

5.2 Minimizing Maximum Delay

The problem of minimizing maximum delay can be formulated as follows:

$$\begin{aligned} \mathcal{M}3: \textit{Minimize} \quad & D_{\max} \\ \textit{Subject to} \quad & D_i(\mathbf{f}) \le D_{\max},\ 1 \le i \le s \\ & L_i \le f_i \le U_i,\ 1 \le i \le n \\ & D_{\max} > 0 \end{aligned}$$

As in NWSA-db, we use the Lagrangian relaxation technique to solve $\mathcal{M}3$. The following is the LRS associated with Lagrange multipliers $\lambda_1, \lambda_2, \ldots, \lambda_s$.

$$\begin{aligned} \mathcal{M}3': \textit{Minimize} \quad & D'(\mathbf{f}) = D_{\max} + \sum_{i=1}^{s} \lambda_i (D_i(\mathbf{f}) - D_{\max}) \\ \textit{Subject to} \quad & L_i \le f_i \le U_i,\ 1 \le i \le n. \end{aligned}$$

Differentiating the objective function with respect to $D_{\max}$ and setting the result equal to 0, we get $\sum_{i=1}^{s} \lambda_i = 1$. Therefore, by Kuhn-Tucker theorem [10], it suffices to use Lagrange multipliers $\lambda_1, \lambda_2, \ldots, \lambda_s$ where $\sum_{i=1}^{s} \lambda_i = 1$. Note that the objective function in $\mathcal{M}3'$ becomes $\sum_{i=1}^{s} \lambda_i D_i(\mathbf{f})$ when $\sum_{i=1}^{s} \lambda_i = 1$. Thus we have the following theorem.

Theorem 4 *Solving the LRS $\mathcal{M}3'$ is equivalent to solving $\mathcal{M}1$.* $\square$

Figure 5 shows the algorithm NWSA-md for solving $\mathcal{M}3$. The sequence $\{\theta_k\}$ in **A5** is the same as that of NWSA-db. Again, it can be shown that NWSA-md converges to an optimal wire-sizing solution (for $\mathcal{M}3$).

Algorithm: NWSA-md
A1. $k = 0$; $f_i = L_i$, $0 \le i \le n$.
A2. $\lambda_i = 1/s$, $1 \le i \le s$.
A3. Call NWSA-wd to solve the LRS associated with $\lambda_1, \ldots, \lambda_s$.
A4. Recursively compute D_i, $1 \le i \le s$; $D_{\max} = \max\{D_i\}$.
A5. $\lambda_i = \lambda_i + \theta_k(D_i - D_{\max}), 1 \le i \le s$. Normalize λ_i such that $\sum_{i=1}^{s} \lambda_i = 1$.
A6. $k = k + 1$.
A7. Repeat **A3**–**A6** until no improvement.

Figure 5: Non-uniform wire-sizing algorithm for minimizing maximum delay

6 Experimental Results

We implemented our algorithms and tested on the five circuits $r1$–$r5$ used in [7] on an IBM RS/6000 workstation. The per micron resistance and capacitance used are $3m\Omega$ and $0.02fF$, respectively. The lower and upper bounds for wire widths are $1\mu m$ and $10\mu m$, respectively. Table 1 lists the names of the circuits, numbers of wire segments in the circuits, delays, runtime, and storage requirements. It shows that NWSA-wd, on the average, reduced the weighted delay by 83.4% after wire-sizing. Furthermore, our algorithm is extremely fast and space-efficient. For example, for the circuit r5 with 6201 wire segments, NWSA-wd needed only about 1.5-second runtime and 1.3-MB storage. In Figure 6, the runtime and storage requirements (represented by the vertical axis) are plotted as a function of the number of wire segments in a circuit (denoted by the horizontal axis). It shows that the runtime and storage requirements of NWSA-wd is linear in the number of wire segments.

Ckt	# Nodes	Delay (ns) Initial	Final	Red%	Time (sec)	Stor (kb)
r1	533	0.744	0.146	80	0.18	148
r2	1195	2.014	0.358	82	0.32	280
r3	1723	3.291	0.543	84	0.45	388
r4	3805	8.644	1.307	85	0.97	812
r5	6201	15.316	2.216	86	1.57	1300
Avg	-	-	-	83.4	-	-

Table 1: Minimizing weighted delay

The experimental results on minimizing area under delay constraints are in Table 2. It show that in order to satisfy the delay constraints (half of the initial delay), NWSA-db only increases average 3.4% of the total area and runtime is only about one minute for r5. Table 3 shows the experimental results for minimizing the maximum sink-delay. It shows that NWSA-db and NWSA-md have similar runtime and

Ckt	# Nodes	Delay (ns)			Area ($10^5 \mu m^2$)			Runtime (sec)	Storage (kbytes)
		Initial	Final	Reduce%	Initial	Final	Increase%		
r1	533	0.775	0.388	50	1.25	1.36	8	4.95	148
r2	1195	2.108	1.054	50	2.48	2.63	3	10.60	280
r3	1723	3.376	1.688	50	3.19	3.31	4	15.15	388
r4	3805	9.087	4.544	50	6.50	6.65	2	35.25	812
r5	6201	15.861	7.931	50	9.73	9.91	2	64.48	1300
Avg	-	-	-	50	-	-	3.8	-	-

Table 2: Minimizing area subject to delay bounds

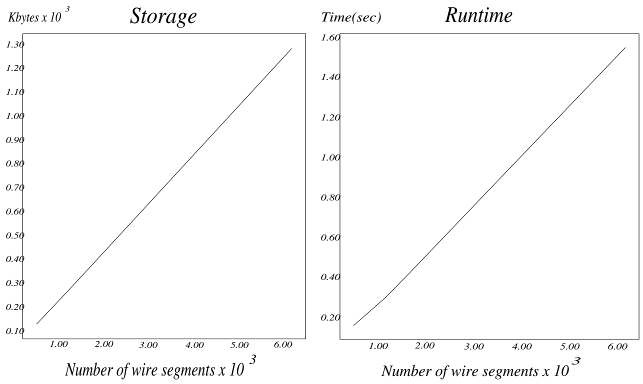

Figure 6: Storage and runtime requirements

storage requirements.

Ckt	# Nodes	Delay (ns)			Time (sec)	Stor (kb)
		Initial	Final	Red%		
r1	533	0.775	0.161	79	3.50	148
r2	1195	2.108	0.379	82	13.38	280
r3	1723	3.376	0.572	83	17.25	388
r4	3805	9.087	1.376	86	54.87	812
r5	6201	15.861	2.312	85	67.04	1300
Avg	-	-	-	83	-	-

Table 3: Minimizing maximum delay

References

[1] C.-P. Chen, Y.-P. Chen, D. F. Wong. "Optimal Wire-sizing formula under the Elmore delay model", *Proc. ACM/IEEE DAC*, pp. 487-490, 1996.

[2] C.-P. Chen, D. F. Wong. "A fast algorithm for optimal wire-sizing under Elmore delay model" *Proc. IEEE ISCAS*, vol. 4, pp. 412-415, 1996.

[3] J. Cong and K. Leung, "Optimal wiresizing under Elmore delay model," *IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems* 14(3), pp. 321–336, 1995.

[4] W. C. Elmore, "The transient response of damped linear networks with particular regard to wide band amplifiers", *J. Applied Physics*, 19(1), 1948.

[5] N. Menezes, S. Pullela, F. Dartu, and L. T. Pillage, "RC interconnect syntheses-a moment fitting approach", *Proc. IEEE/ACM ICCAD*, 1994.

[6] N. Menezes, R. Baldick, and L. T. Pillage, "A sequential quadratic programming approach to concurrent gate and wire sizing", *Proc. IEEE/ACM ICCAD*, 1995.

[7] R.-S. Tsay, "Exact zero skew," *IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems*, Feb. 1993.

[8] M. L. Fisher, "An applications oriented guide to Lagrangian relaxation," *Interfaces*, 15:2, pp. 10-21, March-April 1985.

[9] J. P. Fishburn and C. A. Schevon, "Shaping a distributed-RC line to minimize Elmore delay", *IEEE Transactions on Circuits and Systems-I: Fundamental Theory and Applications*, Vol. 42, No. 12, pp. 1020-1022, December 1995.

[10] D. G. Luenberger, *Linear and Nonlinear Programming*, Addison-Wesley Pub. Company Inc., 1984.

[11] S. S. Sapatnekar, "RC interconnect optimization under the Elmore delay model," *Proc. IEEE/ACM ICCAD*, pp. 387–391, 1994.

[12] J. Cong and Lei He, "Optimal wire-sizing for interconnects with multiple sources", *Proc. IEEE/ACM ICCAD*, pp. 568-574, 1995.