Technische Universität Ilmenau
Institut für Mathematik

Preprint No. M 09/23

The Repeater Tree Construction Problem

Bartoschek, Christoph; Held, Stephan; Maßberg, Jens; Rautenbach, Dieter; Vygen, Jens

2009

Impressum:
Hrsg.: Leiter des Instituts für Mathematik
Weimarer Straße 25
98693 Ilmenau
Tel.: +49 3677 69 3621
Fax: +49 3677 69 3270
http://www.tu-ilmenau.de/ifm/

ISSN xxxx-xxxx

The Repeater Tree Construction Problem

C. Bartoschek[1], S. Held[1], J. Maßberg[1], D. Rautenbach[2], and J. Vygen[1]

[1] Forschungsinstitut für Diskrete Mathematik,
Universität Bonn, Lennéstr. 2, D-53113 Bonn, Germany,
emails: {bartosch,held,massberg,vygen}@or.uni-bonn.de

[2]Institut für Mathematik,
TU Ilmenau, Postfach 100565, D-98684 Ilmenau, Germany
email: {dieter.rautenbach}@tu-ilmenau.de

Abstract

A tree-like substructure on a computer chip whose task it is to carry a signal from a source circuit to possibly many sink circuits and which consists only of wires and so-called repeater circuits is called a repeater tree. We present a mathematical formulation of the optimization problems related to the construction of such repeater trees. Furthermore, we prove theoretical properties of a simple iterative procedure for these problems which was successfully applied in practice.

Keywords: VLSI design; repeater tree; Steiner tree; minimum spanning tree
AMS subject classification: 05C05, 05C85, 68W25, 68W35

1 Introduction

During every computation cycle of a modern highly complex computer chip millions of signals have to travel between circuits at different locations on the chip area. While for most of these signals the distances are relatively small and can be bridged by a pure metal connection between the circuits, there are still many signals which have to travel a relatively long distance. Elementary physical considerations [5] imply that the delay of an electrical signal propagating along a metal connection approximately grows quadratically with the traversed distance. Traditionally, the circuit delay dominated the wire delay and this quadratic growth did not represent a problem. Nowadays though, due to the continuous shrinking of feature sizes [4, 10], an ever growing part of the total delay is caused by wires, and long metal connections have to be split into several parts by inserting so-called *repeaters*. These repeaters just evaluate the boolean identity function and serve no logical purpose within the computation of the chip. Their task is only to linearize the delay as a function of the distance. It is estimated [11] that for the upcoming 45nm and 32nm technologies up to 35% and 70%, respectively, of all circuits on a chip might have to be repeaters.

A tree-like substructure on a chip whose task it is to carry a signal from a source circuit to possibly many sink circuits and which consists only of wires and repeaters is called a *repeater tree*. In [2, 3] we proposed algorithms for the construction of repeater tree topologies and for the actual insertion of repeater circuits into these topologies. During this research we conceived a simple yet relatively accurate delay model which allows a concise mathematical

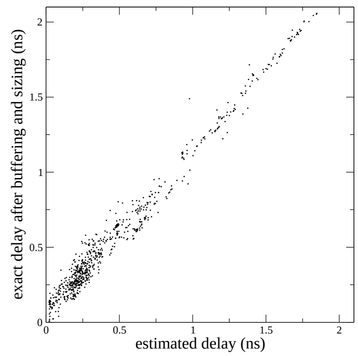

Figure 1: Quality of the delay model

formulation of the repeater tree problem. The purpose of the present paper is to present this formulation, to explain the main optimization goals, and to prove some theoretical properties of the algorithms in [2, 3].

2 The Repeater Tree Problem

An instance of the *repeater tree (topology) problem* consists of

- a *source* $r \in \mathbb{R}^2$,
- a finite non-empty set $S \subseteq \mathbb{R}^2$ of *sinks*,
- a *required arrival time* $a_s \in \mathbb{R}$ for every sink $s \in S$, and
- two numbers $c, d \in \mathbb{R}_{>0}$.

A feasible solution of such an instance is

- a rooted tree $T = (V(T), E(T))$ with vertex set $\{r\} \cup S \cup I$ where $I \subseteq \mathbb{R}^2$ is a set of $|S| - 1$ points such that r is the root of T and has exactly one child, the elements of I are the internal vertices of T and have exactly two children each, and the elements of S are the leaves of T.

In [2, 3] such a feasible solution was called a *repeater tree topology*, because the number, types, and positions of the actual repeaters are not yet determined.

The optimization goals for a repeater tree are related to the wiring, to the number of repeater circuits, and to the timing. We assume that every edge $e = (u, v) \in E(T)$ of T is realized along a path between the two points u and v in the plane which is shortest with respect to some norm $||\cdot||$ on $\mathbb{R}^2$. Furthermore, we assume that repeaters are inserted in a relatively uniform way into all wires in order to linearize the delay within the repeater tree. Hence the wiring and also the number of repeater circuits needed for the physical realization of the edge e are proportional to $||u - v||$. For the entire repeater tree topology, this result in a total cost of

$$l(T) := \sum_{(u,v) \in E(T)} ||u - v||.$$

The delay of the signal starting at the root and travelling through T to the sinks has two components. Let $E[r,s]$ denote the set of edges on the path P in T between the root r and some sink $s \in S$. The linearized delay along the edges of P is modelled by

$$\sum_{(u,v)\in E[r,s]} d||u-v||.$$

Furthermore, every internal vertex on P corresponds to a bifurcation which causes an additive delay of c along P. For the entire path P, these additional delays sum up to

$$c(|E[r,s]|-1).$$

In practice there is sometimes a certain degree of freedom how to distribute the additional delay caused by a bifurcation to the two branches [9].

Altogether, we estimate the delay of the signal along P by the sum of these two components.

Assuming that the signal starts at time 0 at the root, the *slack* at some sink $s \in S$ in T is estimated by

$$\sigma(T,s) := a_s - \sum_{(u,v)\in E[r,s]} d||u-v|| - c(|E[r,s]|-1)$$

and the *worst slack* equals

$$\sigma(T) := \min\{\sigma(T,s) \mid s \in S\}.$$

The restrictions on the number of children of the root and the internal vertices of T imply that the number of sinks contributes logarithmically to the delay, which corresponds to physical experience. The accuracy of our simple delay estimation is shown in Figure 1, which compares our estimation with the real physical delay once the repeater tree has been realized and optimized. The parameters c and d are technology-dependent. For the 65nm technology their values are about $c = 20$ps and $d = 220$ps/mm.

In principle, a repeater tree topology is acceptable with respect to timing if $\sigma(T)$ is non-negative, i.e. the signal arrives at every sink $s \in S$ not later than a_s. Nevertheless, in order to account for inaccurate estimations and manufacturing variation, the worst slack $\sigma(T)$ should have at least some reasonable positive value $\sigma_{\min}$ or should even be maximized.

We can formulate three main optimization scenarios: Determine T such that

(O1) $\sigma(T)$ is maximized, or

(O2) $l(T)$ is minimized, or

(O3) for suitable constants $\alpha, \beta, \sigma_{\min} > 0$, the expression

$$\alpha \min\{\sigma(T), \sigma_{\min}\} - \beta l(T)$$

is maximized.

While scenario (O1) is reasonable for instances which are very timing critical, scenario (O2) is reasonable for very timing uncritical instances. Scenario (O3) is probably the practically most relevant one. In the next section, we will show that (O1) can be solved exactly in polynomial time. In contrast to that, (O2) is hard even for restricted choices of the norm such as the l_1-norm, since it is essentially the Steiner tree problem [6].

3 A Simple Procedure and its Properties

In [2, 3] we considered the following very simple procedure for the construction of repeater tree topologies.

Choose a sink $s_1 \in S$;
$V(T_1) \leftarrow \{r, s_1\}$;
$E(T_1) \leftarrow \{(r, s_1)\}$;
$T_1 \leftarrow (V(T_1), E(T_1))$;
$n \leftarrow |S|$;
for $i = 2$ **to** n **do**
    Choose a sink $s_i \in S \setminus \{s_1, s_2, \ldots, s_{i-1}\}$, an edge $e_i = (u, v) \in E(T_{i-1})$, and an internal vertex $x_i \in \mathbb{R}^2$;
    $V(T_i) \leftarrow V(T_{i-1}) \dot\cup \{x_i\} \dot\cup \{s_i\}$;
    $E(T_i) \leftarrow (E(T_{i-1}) \setminus \{(u, v)\}) \cup \{(u, x_i), (x_i, v), (x_i, s_i)\}$;
    $T_i \leftarrow (V(T_i), E(T_i))$;
end

The procedure inserts the sinks one by one according to some order $s_1, s_2, \ldots, s_n$ starting with a tree containing only the root r and the first sink s_1. The sinks s_i for $i \geq 2$ are inserted by subdividing an edge e_i with a new internal vertex x_i and connecting x_i to s_i. The behaviour of the procedure clearly depends on the choice of the order, the choice of the edge e_i, and the choice of the point $x_i \in \mathbb{R}^2$.

In view of the large number of instances which have to be solved in an acceptable time [2, 3] the simplicity of the above procedure is an important advantage for its practical application. Furthermore, implementing suitable rules for the choice of s_i, e_i, and x_i allows to pursue and balance various practical optimization goals.

We present two variants (P1) and (P2) of the procedure corresponding to the above optimization scenarios (O1) and (O2), respectively.

(P1) The sinks are inserted in an order of non-increasing criticality, where the *criticality* of a sink $s \in S$ is quantified by

$$-(a_s - d||r - s||).$$

(Note that this is the estimated worst slack of a repeater tree topology containing only the one sink s. Since a sink s can be critical because its required arrival time a_s is small and/or because its distance $||r - s||$ to the root is large, this is a reasonable measure for its criticality.)

During the i-th execution of the **for**-loop, the new internal vertex x_i is always chosen at the same position as r — formally this turns $V(T_i)$ into a multiset — and the edge e_i is chosen such that $\sigma(T_i)$ is maximized.

(P2) s_1 is chosen such that $||r - s_1|| = \min\{||r - s|| \mid s \in S\}$ and during the i-th execution of the **for**-loop, s_i, $e_i = (u, v)$, and x_i are chosen such that

$$l(T_i) = l(T_{i-1}) + ||u - x_i|| + ||x_i - v|| + ||x_i - s_i|| - ||u - v||$$

is minimized.

Theorem 1 *The largest achievable worst slack σ_{opt} equals*

$$\sigma^*(S) := \max\left\{\sigma \in \mathbb{R} \mid \sum_{s\in S} 2^{-\left\lfloor \frac{1}{c}(a_s - d||r-s|| - \sigma)\right\rfloor} \le 1\right\},$$

and (P1) generates a repeater tree topology $T_{(P1)}$ with $\sigma\left(T_{(P1)}\right) = \sigma_{\text{opt}}$.

Proof: Let $a'_s = a_s - d||r-s||$ for $s \in S$. Let T be an arbitrary repeater tree topology. By the definition of $\sigma(T)$ and the triangle-inequality for $||\cdot||$, we obtain

$$|E[r,s]| - 1 \quad \le \quad \left\lfloor \frac{1}{c}\left(a_s - \sum_{(u,v)\in E[r,s]} d||u-v|| - \sigma(T)\right)\right\rfloor \le \left\lfloor \frac{1}{c}\left(a'_s - \sigma(T)\right)\right\rfloor$$

for every $s \in S$. Since the unique child of the root r is itself the root of a binary subtree of T in which each sink $s \in S$ has depth exactly $|E[r,s]| - 1$, Kraft's inequality [8] implies

$$\sum_{s\in S} 2^{-\left\lfloor \frac{1}{c}(a'_s - \sigma(T))\right\rfloor} \quad \le \quad \sum_{s\in S} 2^{-|E[r,s]|+1} \le 1.$$

By the definition of $\sigma^*(S)$, this implies $\sigma(T) \le \sigma^*(S)$. Since T was arbitrary, we obtain $\sigma_{\text{opt}} \le \sigma^*(S)$.

It remains to prove that $\sigma\left(T_{(P1)}\right) = \sigma_{\text{opt}} = \sigma^*(S)$, which we will do by induction on $n = |S|$. For $n = 1$, the statement is trivial. Now let $n \ge 2$. Let s_n be the last sink inserted by (P1), i.e. $a'_{s_n} = \max\{a'_s \mid s \in S\}$. Let $S' = S \setminus \{s_n\}$.

Claim

$$\text{frac}\left(\frac{\sigma^*(S)}{c}\right) \quad \in \quad \left\{\text{frac}\left(\frac{a'_s}{c}\right) \mid s \in S'\right\} \tag{1}$$

where $\text{frac}(x) := x - \lfloor x \rfloor$ denotes the fractional part of $x \in \mathbb{R}$.

Proof of the claim: Note that the definition of $\sigma^*(S)$ implies that $\frac{1}{c}\left(a'_s - \sigma^*(S)\right)$ is an integer for at least one $s \in S$. If the claim is false, then $\frac{1}{c}\left(a'_{s_n} - \sigma^*(S)\right) \in \mathbb{Z}$ and $\frac{1}{c}\left(a'_s - \sigma^*(S)\right) \notin \mathbb{Z}$ for every $s \in S'$. Since $a'_{s_n} - \sigma^*(S) \ge a'_s - \sigma^*(S)$ for every $s \in S'$, this implies

$$\left\lfloor \frac{1}{c}\left(a'_{s_n} - \sigma^*(S)\right)\right\rfloor > \max\left\{\left\lfloor \frac{1}{c}\left(a'_s - \sigma^*(S)\right)\right\rfloor \mid s \in S'\right\}$$

and hence

$$\sum_{s\in S} 2^{-\left\lfloor \frac{1}{c}(a'_s - \sigma^*(S))\right\rfloor} \le 1 - 2^{-\left\lfloor \frac{1}{c}\left(a'_{s_n} - \sigma^*(S)\right)\right\rfloor}.$$

Now, for some sufficiently small $\epsilon > 0$, we obtain

$$\sum_{s\in S} 2^{-\left\lfloor \frac{1}{c}(a'_s - (\sigma^*(S)+\epsilon))\right\rfloor} = 2^{-\left\lfloor \frac{1}{c}\left(a'_{s_n} - \sigma^*(S)\right)\right\rfloor + 1} + \sum_{s\in S'} 2^{-\left\lfloor \frac{1}{c}(a'_s - \sigma^*(S))\right\rfloor} \le 1$$

which contradicts the definition of $\sigma^*(S)$ and completes the proof of the claim. $\square$

Let $T'_{(P1)}$ denote the tree produced by (P1) just before the insertion of the last sink s_n. By induction, $\sigma\left(T'_{(P1)}\right) = \sigma^*(S')$.

First, we assume that there is some sink $s' \in S'$ such that within $T'_{(P1)}$

$$|E[r, s']| - 1 < \left\lfloor \frac{1}{c}\left(a'_{s'} - \sigma^*(S')\right) \right\rfloor .$$

Choosing e_n as the edge of $T'_{(P1)}$ leading to s', results in a tree T such that

$$\sigma^*(S) \geq \sigma_{\text{opt}} \geq \sigma\left(T_{(P1)}\right) \geq \sigma(T) = \sigma^*(S') \geq \sigma^*(S),$$

which implies $\sigma\left(T_{(P1)}\right) = \sigma_{\text{opt}} = \sigma^*(S)$.

Next, we assume that within $T'_{(P1)}$

$$|E[r, s]| - 1 = \left\lfloor \frac{1}{c}\left(a'_{s} - \sigma^*(S')\right) \right\rfloor$$

for every $s \in S'$. This implies

$$\sum_{s\in S} 2^{-\left\lfloor \frac{1}{c}\left(a'_s - \sigma^*(S')\right)\right\rfloor} > \sum_{s\in S'} 2^{-\left\lfloor \frac{1}{c}\left(a'_s - \sigma^*(S')\right)\right\rfloor} = 1$$

and hence $\sigma^*(S) < \sigma^*(S')$. By (1), we obtain

$$\begin{aligned}
\sigma^*(S) &\leq \max\left\{\sigma \mid \sigma < \sigma^*(S'),\, \text{frac}\left(\frac{\sigma}{c}\right) \in \left\{\text{frac}\left(\frac{a'_s}{c}\right) \mid s \in S'\right\}\right\} \\
&= \max\left\{\sigma \mid \sigma < \sigma^*(S'),\, \text{frac}\left(\frac{\sigma - \sigma^*(S')}{c}\right) \in \left\{\text{frac}\left(\frac{a'_s - \sigma^*(S')}{c}\right) \mid s \in S'\right\}\right\} \\
&= c\max\left\{x \mid x < \frac{\sigma^*(S')}{c},\, \text{frac}\left(x - \frac{\sigma^*(S')}{c}\right) \in \left\{\text{frac}\left(\frac{a'_s - \sigma^*(S')}{c}\right) \mid s \in S'\right\}\right\} \\
&= c\left(\frac{\sigma^*(S')}{c} - 1 + \max\left\{\text{frac}\left(\frac{a'_s - \sigma^*(S')}{c}\right) \mid s \in S'\right\}\right) \\
&= \sigma^*(S') - c(1-\delta)
\end{aligned}$$

for

$$\delta = \max\left\{\text{frac}\left(\frac{a'_s - \sigma^*(S')}{c}\right) \mid s \in S'\right\}.$$

If $s' \in S'$ is such that

$$\delta = \text{frac}\left(\frac{a'_{s'} - \sigma^*(S')}{c}\right),$$

then choosing e_n as the edge of $T'_{(P1)}$ leading to s', results in a tree T such that

$$\sigma^*(S) \geq \sigma_{\text{opt}} \geq \sigma\left(T_{(P1)}\right) \geq \sigma(T) = \sigma^*(S') - c(1-\delta) \geq \sigma^*(S),$$

which implies $\sigma\left(T_{(P1)}\right) = \sigma_{\text{opt}} = \sigma^*(S)$ and completes the proof. $\square$

Theorem 2 *(P2) generates a repeater tree topology T for which $l(T)$ is at most the total length of a minimum spanning tree on $\{r\} \cup S$ with respect to $||\cdot||$.*

Proof: Let $n = |S|$ and for $i = 0, 1, \ldots, n$, let $\overline{T}_i$ denote the forest which is the union of the tree produced by (P2) after the insertion of the first i sinks and the remaining $n - i$ sinks as isolated vertices. Note that $\overline{T}_0$ has vertex set $\{r\} \cup S$ and no edge, while for $1 \le i \le n$, $\overline{T}_i$ has vertex set $\{r\} \cup S \cup \{x_j \mid 2 \le j \le i\}$ and $2i - 1$ edges.

Let $F_0 = (V(F_0), E(F_0))$ be a spanning tree on $V(F_0) = \{r\} \cup S$ such that

$$l(F_0) = \sum_{uv \in E(F_0)} ||u - v||$$

is minimum. For $i = 1, 2, \ldots, n$, let $F_i = (V(F_i), E(F_i))$ arise from

$$\left(V\left(\overline{T}_i\right), E(F_{i-1}) \cup E\left(\overline{T}_i\right)\right)$$

by deleting an edge $e \in E(F_{i-1}) \cap E(F_0)$ which has exactly one endvertex in $V\left(T_{i-1}\right)$ such that F_i is a tree. (Note that this uniquely determines F_i.)

Since (P2) has the freedom to use the edges of F_0, the specification of the insertion order and the locations of the internal vertices in (P2) imply that

$$l(F_0) \ge l(F_1) \ge l(F_2) \ge \ldots \ge l(F_n).$$

Since $F_n = T_n$ the proof is complete. $\Box$

For the l_1-norm, the well-known result of Hwang [7] together with Theorem 2 imply that (P2) is an approximation algorithm for the l_1-minimum Steiner tree on the set $\{r\} \cup S$ with approximation guarantee $3/2$.

We have seen in Theorems 1 and 2 that different insertion orders are favourable for different optimization scenarios such as (O1) and (O2).

Alon and Azar [1] gave an example showing that for the online rectilinear Steiner tree problem the best approximation ratio we can achieve is $\Theta(\log n / \log\log n)$, where n is the number of terminals. Hence inserting the sinks in an order disregarding the locations, like in (P1), can lead to long Steiner trees, no matter how we decide where to insert the sinks.

The next example shows that inserting the sinks in an order different from the one considered in (P1) but still choosing the edge e_i as in (P1) results in a repeater tree topology whose worst slack can be much smaller than the largest achievable worst slack.

Example 3 Let $c = 1$, $d = 0$ and $a \in \mathbb{N}$. We consider the following sequences of $-a$'s and 0's

$$\begin{aligned}
A(1) &= (-a, 0), \\
A(2) &= (A(1), -a, 0), \\
A(3) &= (A(2), -a, \underbrace{0, \ldots\ldots, 0}_{1+(2^1-1)(a+2)}), \\
A(4) &= (A(3), -a, \underbrace{0, \ldots\ldots\ldots, 0}_{1+(2^2-1)(a+2)}), \ldots,
\end{aligned}$$

i.e. for $l \ge 2$, the sequence $A(l)$ is the concatenation of $A(l-1)$, one $-a$, and a sequence of 0's of length $1 + \left(2^{l-2} - 1\right)(a + 2)$.

If the entries of $A(l)$ are considered as the requires arrival times of an instance of the repeater tree topology problem, then Theorem 1 together with the choice of c and d imply that the largest achievable worst slack for this instance equals

$$\left\lfloor -\log_2\left(l2^a + \left(1+\sum_{i=2}^{l}\left(1+(2^{i-2}-1)(a+2)\right)\right)2^0\right)\right\rfloor.$$

For $l = a+1$ this is at least $-2-a-\log_2(a+2)$.

If we insert the sinks in the order as specified by the sequences $A(l)$, and always choose the edge into which we insert the next internal vertex such that the worst slack is maximized, then the following sequence of topologies can arise: $T(1)$ is the topology with two exactly sinks at depth 2. The worst slack of $T(1)$ is $-(a+2)$. For $l \geq 2$, $T(l)$ arises from $T(l-1)$ by (a) subdividing the edge of $T(l-1)$ incident with the root with a new vertex x, (b) appending an edge (x, y) to x, (c) attaching to y a complete binary tree B of depth $l-2$, (d) attaching to one leaf of B two new leaves corresponding to sinks with required arrival times $-a$ and 0, and (e) attaching to each of the remaining $2^{l-2}-1$ many leaves of B a binary tree Δ which has $a+2$ leaves, all corresponding to sinks of arrival times 0, whose depths in Δ are $1, 2, 3, \ldots, a-1, a, a+1, a+1$. Note that this uniquely determines $T(l)$.

Clearly, the worst slack in $T(l)$ equals $-a-(l+1)$. Hence for $l = a+1$, the worst slack equals $-2a-2$, which differs approximately by a factor of 2 from the largest achievable worst slack as calculated above.

This example, however, does not show that there is no online algorithm for approximately maximizing the worst slack, say up to an additive constant of c. It is an open question to find a bicriteria approximation algorithm, or an algorithm for (O3).

References

[1] N. Alon and Y. Azar, On-line Steiner trees in the Euclidean plane, *Discrete and Computational Geometry* **10** (1993), 113–121.

[2] C. Bartoschek, S. Held, D. Rautenbach, and J. Vygen, Efficient generation of short and fast repeater tree topologies, in: Proceedings of the International Symposium on Physical Design (2006), 120–127.

[3] C. Bartoschek, S. Held, D. Rautenbach, and J. Vygen, Fast buffering for optimizing worst slack and resource consumption in repeater trees, in: Proceedings of the International Symposium on Physical Design (2009), 43–50.

[4] J. Cong, An interconnect-centric design flow for nanometer technologies, in: Proceedings of the IEEE **89** (2001), 505–528.

[5] W.C. Elmore, The transient response of damped linear networks with particular regard to wideband amplifiers, *Journal of Applied Physics* **19** (1948), 55–63.

[6] M.R. Garey, and D.S. Johnson, The rectilinear Steiner tree problem is NP-complete, *SIAM Journal on Applied Mathematics* **32** (1977), 826–834.

[7] F.K. Hwang, On steiner minimal trees with rectilinear distance, *SIAM Journal of Applied Mathematics* **30** (1976), 104–114.

[8] L.G. Kraft, A device for quantizing grouping and coding amplitude modulated pulses, Master thesis, EE Dept., MIT, Cambridge 1949.

[9] J. Maßberg and D. Rautenbach, Binary trees with choosable edge lengths, to appear in *Information Processing Letters*.

[10] G.E. Moore, Cramming more components onto integrated circuits, *Electronics* **38** (1965), 114–117.

[11] P. Saxena, N. Menezes, P. Cocchini, and D. Kirkpatrick, The scaling challenge: can correct-by-construction design help?, in: Proceedings of the International Symposium on Physical Design (2003), 51–58.