Congestion Prediction in Early Stages of Physical Design

CHIU-WING SHAM
The Hong Kong Polytechnic University
EVANGELINE F. Y. YOUNG
The Chinese University of Hong Kong
and
JINGWEI LU
The Hong Kong Polytechnic University

Routability optimization has become a major concern in physical design of VLSI circuits. Due to the recent advances in VLSI technology, interconnect has become a dominant factor of the overall performance of a circuit. In order to optimize interconnect cost, we need a good congestion estimation method to predict routability in the early designing stages. Many congestion models have been proposed but there's still a lot of room for improvement. Besides, routers will perform rip-up and reroute operations to prevent overflow, but most models do not consider this case. The outcome is that the existing models will usually underestimate the routability. In this paper, we have a comprehensive study on our proposed congestion models. Results show that the estimation results of our approaches are always more accurate than the previous congestion models.

Categories and Subject Descriptors: B.7.2 [**Integrated Circuit**]: Design Aids

General Terms: Performance

Additional Key Words and Phrases: Estimation, placement, floorplanning

ACM Reference Format:

Sham, C.-W., Young, E. F. Y., and Lu, J. 2009. Congestion prediction in early stages of physical design. ACM Trans. Des. Autom. Elect. Syst., 14, 1, Article 12 (January 2009), 18 pages, DOI = 10.1145/1455229.1455241 http://doi.acm.org/10.1145/1455229.1455241

The work described in this article was partially supported by the Competitive Earmarked Research Grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. PolyU 5262/07E).
Author's address: C.-W. Sham; email: cwsham@alumni.cuhk.net.

1. INTRODUCTION

1.1 Motivations

The routability problem is a demand and supply problem of the routing resources. In the early stages of the design cycle, the shapes and locations of the modules on a chip are planned, and its result will greatly affect the overall performance of the final design. In some advanced systems using the deep submicron technology today, the extremely high design densities will result in a major escalation in routing demand. Overcongestion will deteriorate circuit performance or even lead to an unroutable solution. Thus, routability optimization has become a major concern in physical design. Unfortunately, minimizing total wirelength does not have significant impact on routability [Wang et al. 2000]. We need an accurate congestion prediction and an efficient congestion removal technique.

Excessive congestion will result in a local shortage of routing resources. This will lead to a large expansion in area, or even an unroutable design failing to achieve timing closure after detailed routing. In this case, the design process must be restarted from an early stage such as floorplanning and placement. Thus, it is desirable to detect and remove congested regions in the early designing stages. However, in an automated IC implementation flow, congestion information will be available only after detailed routing. A good congestion model is needed for accurate interconnect analysis and prediction during the early stages of the design process.

1.2 Related Works

Because of the importance of this congestion estimation problem, many models have been proposed. In some papers [Chen et al. 1999; Chang et al. 2000; Ma et al. 2003], a packing is divided into tiles and congestion is estimated in each tile, assuming that each net is routed in either L- or Z-shape. In Lai et al. [2003], the congestion model used is the average net density on the boundaries of different regions in a floorplan. In other papers [Kusnadi and Carothers 1999; Lou et al. 2001], probabilistic analysis is performed to estimate congestion and routability. They assume that all feasible routes have the same probability of being selected. In practice, routes of less bends are more desirable. In Kahng and Xu [2003] and Westra et al. [2004], extended versions of Lou et al. [2001] are proposed. The authors take into account the impact of the number of bends in a routing path on the probability of occurrence of the path. However, the accuracies of their congestion models will depend on the accuracies of their predictions on the distribution of the number of bends in the routed circuit. Yang et al. [2002] proposed to predict congestion by using the Rent's rule. However, connections of the nets are already known in the floorplanning and placement stage, and we should be able to predict congestion more accurately than simply using the Rent's rule. Wang and Sarrafzadeh [2000] and Wang et al. [2000] proposed to use global routers to estimate congestion, which will be more accurate but the runtime penalty is high. Saeedi et al. [2006] presented a probabilistic congestion prediction method based on router's intelligence. Taghavi et al. [2007] gives an tutorial on all recent congestion technique and show the importance of the congestion prediction.

1.3 Our Contributions

Congestion prediction is an important part of interconnect planning in the early stages of the physical design cycle. Although some congestion models have been proposed, the accuracies of the predictions still have a lot of room for improvement. In this article, we have a comprehensive study different congestion models: Lou's Model [Lou et al. 2001], Westra's Model [Westra et al. 2004], 3-step approach [Sham and Young 2005a], SMD, and Detour Model [Sham and Young 2005b]. Results show that the estimation results of the congestion are the trade-off between the accuracy and the runtime. In addition, the 3-step approach is the most efficient one to be used with considering the runtime and accuracy.

This article is organized as follows. First, an analysis of net bending will be given in Section 4. Details of SMD model, Detour model (with estimation of detoured length), and the 3-Step approach (extension of SMD model) will be described in section 5, 6, and 7 respectively. We have also considered blockages, which will be discussed in Section 8. Finally, the experimental results will be shown in Section 9.

2. TERMINOLOGY

The notations used in this article are shown in Table I.

3. PROBLEM FORMULATION

Congestion modeling is an important part of interconnect estimation during the floorplanning and placement stage. Given a packing which is partitioned into $l_h \times l_w$ tiles (according to the length of tile, t_l), we will calculate the net density at each tile according to the congestion model. We can obtain the congestion information from the net densities and evaluate the routability of the packing. Notice that the multi-pin nets are broken down into 2-pin nets first before congestion estimation.

4. INVESTIGATION OF NET BENDINGS

In real routing, the routes of less bends are more preferable, so the feasible routes with less bends should have a higher probability of being selected. We cannot assumes that all feasible routes have the same probability of being selected. Lou's Model [Lou et al. 2001] has made this assumption and so it yields a high probabilistic usage at the center of the bounding box spanned by a net.

On the other hands, some previous congestion models [Kahng and Xu 2003; Westra et al. 2004] may make use of the information of net bendings. They claim that the distribution of bends of each net are similar or tools-dependent only and they calculate the probabilistic usage of each net at each tile accordingly. It means that the accuracy of their congestion estimation may depend on the accuracy of the distribution of bends.

Table I. Notations Used in This Article

Notation	Description
t_l	Length of a tile
c^h_{max}	Maximum horizontal wire capacity inside a tile
c^v_{max}	Maximum vertical wire capacity inside a tile
$c_k(r)$	The number of tiles that is r tiles from the source of net k
DT_k	The shortest Manhattan distance between the source and sink of net k
$d_k(x, y)$	The distance from the source of net k to tile (x, y)
$P_k(x, y)$	A rough estimation of the probability of net k passing through tile (x, y)
$P(x, y)$	Congestion at tile (x, y) obtained from the preliminary estimation step
$W(x, y)$	The weight of tile (x, y)
CF_k	Congestion factor of the net k
l^k_d	detoured length of the net k
$E_k(x, y)$	The probability of net k passing through (x, y)
$E^h_k(x, y)$	The probability of net k passing through (x, y) horizontally
$E^v_k(x, y)$	The probability of net k passing through (x, y) vertically
$E^h(x, y)$	The expected number of wires passing through (x, y) horizontally
$E^v(x, y)$	The expected number of wires passing through (x, y) vertically
$A^h(x, y)$	The actual number of wires passing through (x, y) horizontally obtained from the global router
$A^v(x, y)$	The actual number of wires passing through (x, y) vertically obtained from the global router
(s^x_k, s^y_k)	Co-ordinate of the source tile of net k
(t^x_k, t^y_k)	Co-ordinate of the sink tile of net k
T^d_k	The set of extra tiles when the detoured nets pass through outside the bounding box of net k
T_k	The set of tiles inside the bounding box of net k
$T_k(d)$	The set of tiles inside the bounding box of net k and being d tiles away from the source
$B(x, y)$	Degree of blocking at tile (x, y)

Table II. The Information of Net Bending under Different Routing Environment

Test Cases		Number of Bends (%)				Test Cases		Number of Bends (%)			
		1	2	3	4+			1	2	3	4+
hp	$R1$	32.29	36.43	13.53	17.01	*ami*49	$R1$	46.48	27.51	9.24	13.29
	$R2$	34.90	7.26	10.25	7.39		$R2$	49.20	38.06	6.75	5.99
	$R3$	38.15	54.09	3.31	3.63		$R3$	51.05	44.75	2.34	1.86
apte	$R1$	43.06	25.96	13.55	17.10	*playout*	$R1$	41.85	54.13	3.48	0.49
	$R2$	48.03	30.87	12.40	8.63		$R2$	41.93	55.00	2.96	0.11
	$R3$	52.95	33.83	6.61	6.61		$R3$	41.96	56.48	1.56	0.00
*ami*33	$R1$	35.57	46.20	11.77	6.39						
	$R2$	38.79	51.34	7.11	2.75						
	$R3$	41.80	56.33	1.11	0.75						

However, we have investigated that the distribution of bends may also depend on the congestion of the packing. The statistics are shown in Table II. In the experiments, the test cases used are MCNC benchmark circuits *hp*, *apte*, *ami*33, *ami*49, and *playout*. The detailed information of the testing circuits are shown in Table III. Notice that the number of 2-pin nets are obtained after net decomposition. We will perform global routing using a maze router on each packing under different routing environment $R1$, $R2$ and $R3$. The maximum

Table III. Information of the Test Cases

Test Cases	No. of Cells	No. of Nets	No. of 2-pin nets	t_l (μm)	Max. Wire Capacity in a Tile		
					Rl	$R2$	$R3$
hp	11	83	157	3	7.0	10.0	15.0
apte	9	97	183	4	7.0	10.0	15.0
$ami33$	33	123	305	25	6.0	8.0	12.0
$ami49$	49	408	526	150	7.0	10.0	15.0
playout	62	1611	2138	30	150.0	180.0	240.0

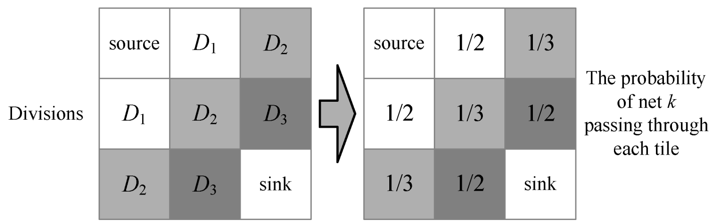

Fig. 1. SMD model for a two-pin net.

wire capacities of a tile will be different under different routing environments. We find that the distribution of bends under different maximum wire capacities is not similar. We can see that if a packing has a smaller maximum wire capacity, the packing should be more congested and so more nets may have more bends. It means that if we want to predict the congestion of the packing accurately, we need to obtain the distribution of bends properly under different situations.

This investigation shows that the number of net bendings depends on the congestion of the circuit. If we need to estimate the number of net bendings correctly, we need to have an accurate congestion prediction first. Hence, if congestion prediction takes this net bending issue into account, it becomes a deadlock. It is the reason that we do not consider net bending in the proposed congestion models.

5. SMD MODEL

When we assume that all the nets are routed in their shortest Manhattan distances, the tiles within the smallest bounding box of net k can be divided into $DT_k - 1$ divisions where DT_k is the shortest Manhattan distance between the source and sink of net k. An example is shown in Figure 1. The tiles are divided into three divisions D_1, D_2, and D_3. Intuitively, if the nets are restricted to be routed within the bounding box with the shortest Manhattan distance, the nets will pass through exactly one tile in each division of the corresponding smallest bounding box. Instead of assuming that the probability of each possible route is the same, we propose a new congestion model, SMD model, assuming that a net will pass through the tiles in the same division with the similar probability. Thus, all the tiles that have the same distance from the source or sink of a net k will have the same probability of being passed through by net k.

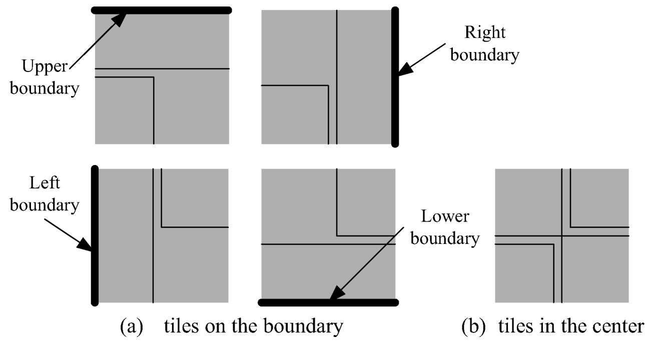

Fig. 2. Possible routes inside a tile (routed from the upper-left corner to the lower-right corner).

Let $s_k(x, y)$ denote the distance from the source of net k to tile (x, y). The tiles having the same distance from the source of net k will be grouped in the same division. Let $c_k(r)$ be the number of tiles that is r tiles from the source of net k. Hence, the probability of net k passing through (x, y), $P_k(x, y)$, can be calculated by the following equation:

$$P_k(x, y) = \frac{1}{c_k(d_k(x, y))}. \tag{1}$$

In addition, a net may pass through a tile either horizontally or vertically. When a net is routed from the upper-left corner to the lower-right corner of the bounding box, the net may pass through a tile with a path as shown in Figure 2. If the tile is on the boundary of the bounding box, the route may pass through the tile in two ways. The four different cases of the tile lying along the top, the left, the bottom and the right boundary are shown in Figure 2a. If the tile is on the left or right (at the top or bottom) of the bounding box, the length of the route passing through the tile horizontally (vertically) is $\frac{0.5t_l}{2}$ and the length of the route passing through the tile vertically (horizontally) is $\frac{1.5t_l}{2}$. If a tile is not on the boundary of the bounding box, the net may pass through the tile in four different ways. They are shown in Figure 2b. In this case, the length of the route passing through the tile horizontally or vertically is $\frac{2t_l}{4}$. Thus, we can calculate $E_k^h(x, y)$ and $E_k^v(x, y)$ by the following equations:

$$E_k^h(x, y) = \begin{cases} \frac{3 \times E_k(x,y)}{4} & : y = s_k^y \text{ or } y = t_k^y \left(x \neq s_k^x \text{ and } x \neq t_k^x\right) \\ \frac{E_k(x,y)}{4} & : x = s_k^x \text{ or } x = t_k^x \left(y \neq s_k^y \text{ and } y \neq t_k^y\right) \\ \frac{E_k(x,y)}{2} & : \text{ otherwise} \end{cases} \tag{2a}$$

$$E_k^v(x, y) = \begin{cases} \frac{E_k(x,y)}{4} & : y = s_k^y \text{ or } y = t_k^y \left(x \neq s_k^x \text{ and } x \neq t_k^x\right) \\ \frac{3 \times E_k(x,y)}{4} & : x = s_k^x \text{ or } x = t_k^x \left(y \neq s_k^y \text{ and } y \neq t_k^y\right) \\ \frac{E_k(x,y)}{2} & : \text{ otherwise.} \end{cases} \tag{2b}$$

Table IV. Percentage of Detoured Nets

Test Cases	Percentage of Detoured Nets (%)		
	$R1$	$R2$	$R3$
hp	11.232	8.664	6.432
apte	17.75	16.363	10.713
*ami*33	8.794	6.432	3.804
*ami*49	15.747	10.633	5.452
playout	0.389	0.256	0.081

Finally, the expected number of wires passing through (x, y) horizontally and vertically, $E^h(x, y)$ ($E^v(x, y)$), can be calculated by the following equations:

$$E^h(x, y) = \sum_{\text{all net } k} E^h_k(x, y) \tag{3a}$$

$$E^v(x, y) = \sum_{\text{all net } k} E^v_k(x, y). \tag{3b}$$

6. DETOUR MODEL

The congestion models discussed before assume that all nets are routed in their shortest Manhattan distances. However, the congestion model can be more accurate when we consider net routing without this assumption. This is reasonable as it is nearly impossible to route all the nets in their shortest Manhattan distances for any large circuit. In this section, we will also propose a new congestion model (called Detour model [Sham and Young 2005b]) in which each net is not necessarily routed in its shortest Manhattan distance. Detour model is proposed based on the SMD model. It assumes that a route may have detours.

6.1 Estimation of Detoured Length

In this section, we show how to estimate the congestion of the tiles when nets may detour. Obviously, the nets may detour only when some locations of the packings are very congested. We perform global routing on the packings (we use the same experiment settings as the experiments in Table II) to illustrate the percentage of detoured nets under different routing environments. The experimental results are shown in Table IV. We can see that if the packing is more congested, more nets may detour. Thus, we can first use SMD model to evaluate the congestion of the packing and calculate the congestion factor of each net k by the following equation:

$$CF_k = \sum_{(x,y)\in T} \frac{2 \times \left(E^h(x, y) - E^h_k(x, y) + E^v(x, y) - E^v_k(x, y)\right)}{\left|s^k_x - t^k_x\right| + 1 \times \left|s^k_y - t^k_y\right| + 1} \times \left(c^h_{max} + c^v_{max}\right), \tag{4}$$

where $s^k_x \le x \le t^k_x$ and $s^k_y \le y \le t^k_y$.

From the equation, we can see that if CF_k is larger than one, it means that the bounding box is very congested for net k and the net k may likely detour.

Table V. Improvement of Wirelength Estimation

Test Cases	Actual wirelength	Estimated wirelength	RMST
hp	1091.10(0.00%)	1079.44(1.56%)	1068.80(2.18%)
apte	1888.24(0.00%)	1873.92(1.01%)	1846.60(2.3%)
*ami*33	2065.78(0.00%)	2054.18(0.61%)	2039.90(1.25%)
*ami*49	3127.44(0.00%)	3028.04(3.17%)	2940.30(5.38%)
playout	12307.02(0.00%)	12305.78(0.06%)	12304.50(0.06%)

Thus, we can estimate the length of detour of net k by the following equation (if l_d^k is smaller than zero, it is adjusted to zero):

$$l_d^k = \lfloor (CF_k - 1) \times DT_k \rfloor \quad (5)$$

We have performed global routing on the packings (we use the same experiment settings as the experiments in Table II) to verify the estimation of detoured length. The experimental results are shown in Table V. From the experiments, we can see the error between our estimated wirelength (the multipin nets are decomposed into 2-pins nets by rectilinear minimum spanning tree first) and the actual wirelength by global routing is always smaller than the error between the wirelength of rectilinear minimum spanning tree (RMST) and the actual wirelength. It means that we can always estimate the wirelength more accurately by using congestion factor.

6.2 Congestion Estimation

As the detoured nets may pass through outside the bounding box, some extra tiles may be passed through by those nets. These extra tiles are the tiles when the detoured nets pass through outside the bounding box of net k with detoured length less than l_d^k. First, we define T_k^d as those extra tiles for a net k. In real cases, $P_k(x, y)$ for $(x, y) \in T_k^d$ will depend on the location of the congestion area. Detours happen because of some random over-congested tiles. Because of this random situation, it is reasonable to assume that all tiles (x, y) in T_k^d have the same probability. Hence, we can build the congestion map based on SMD model. An example is shown in Figure 3.

The detoured net may often pass through outside the bounding box with less bends. If edge shifting [Pan and Chu 2006] is used, the most common case is that either the whole vertical or horizontal path is outside the bounding box. Thus, either all horizontal segments or all vertical segments are passing through outside the bounding box. In this case, we assume that half of the wirelength of the detoured net will be contributed by the tile in T_k^d.

The detoured net may often pass through outside the bounding box with less bends. The most common case is that either the whole vertical or horizontal path is outside the bounding box. Thus, we assume that half of the wirelength of the detoured net will be contributed by the tile in T_k^d. Hence, we can calculate $P_k(x, y)$ where $(x, y) \in T_k^d$ by the following equation:

$$E_k(x, y) = \frac{DT_k + l_d^k}{2 \times |T_k^d|} : (x, y) \in T_k^d, \quad (6)$$

where $|T_k^d|$ is the number of tiles in T_k^d.

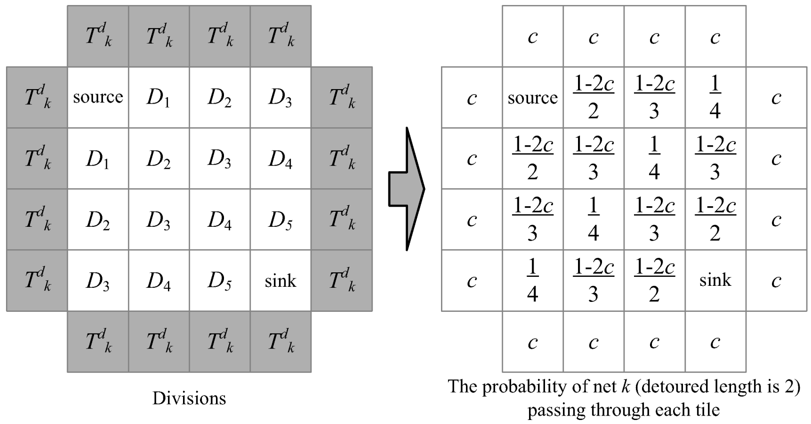

Fig. 3. Detour model for a two-pin net.

Hence, we can calculate $E_k(x, y)$ of other tiles (for any $(x, y) \notin T_k^d$) that inside the bounding box following equation:

$$E_k(x, y) = \frac{1 - c * |T_k^d|_{s_k(x,y)}}{c_k(s_k(x, y))}, \tag{7}$$

where $|T_k^d|_{s_k(x,y)}$ is the number of tiles in same diagonal of the division $s_k(x, y)$ and c is $E_k(x, y)$ for any $(x, y) \in T_k^d$. In addition, a net may pass through a tile either horizontally or vertically. We will calculate $E_k^h(x, y)$ and $E_k^v(x, y)$ by Equation (2a) and (2b) respectively.

7. 3-STEP APPROACH

The estimation process is divided into three steps: preliminary estimation, detailed estimation and congestion redistribution. To avoid overestimating congestion, we perform a preliminary estimation step first to determine which regions are likely to be overcongested. A region should be more attractive to net routing if it is less congested. Then, we will make use of this information to predict the congestion measures during the detailed estimation step. We use a SMD model described in Section 5 because of its simplicity and experimental results have shown that this model can give accurate estimations. Finally, congestion redistribution will be performed to simulate the rip-up and reroute operations of the detailed routing step by moving wires from overcongested regions to less congested regions. We use a 3-step approach [Sham and Young 2005a] as follows:

—Preliminary Estimation: We estimate the congestion measure at each tile roughly according to the bounding box of each net so that we can determine which regions are likely to be overcongested.
—Detailed Estimation: Based on the information obtained from the preliminary estimation step, we estimate the congestion measure at each tile by using a diagonal-based congestion model.

—Congestion Redistribution: We will simulate the rip-up and re-route process of the routing stage by moving wires from overcongested tiles to less congested tiles.

7.1 Preliminary Estimation

In practice, we will choose to route a net over the tiles that are less congested to prevent overflow. It means that some tiles are more attractive to net routing and some tiles are less. However, this fact is usually ignored in traditional congestion models. In our approach, a preliminary estimation of the congestion map will be performed to obtain this information. If a rough estimation of the congestion measure of a tile, $P(x, y)$, is above the maximum wire capacity, the tile (x, y) will be less attractive to net routing. On the other hand, if $P(x, y)$ is well below the maximum wire capacity, the tile (x, y) will be more attractive to net routing. We will make use of these $P(x, y)$ values to improve the accuracy of the detailed estimation step.

In this preliminary estimation step, we assume that all the tiles inside the bounding box of a net k, T_k, have the same probability, $P_k(x, y)$, of being passed through by net k. In addition, we assume that the nets can be routed in their shortest Manhattan distances. The wirelength and the area of the bounding box can be computed as $|t_k^x - s_k^x| + |t_k^y - s_k^y| + 1$ and $(|t_k^x - s_k^x| + 1) \times (|t_k^y - s_k^y| + 1)$ respectively. $P_k(x, y)$ can thus be calculated by the following equation:

$$P_k(x, y) = \frac{|t_k^x - s_k^x| + |t_k^y - s_k^y| + 1}{\left(|t_k^x - s_k^x| + 1\right) \times \left(|t_k^y - s_k^y| + 1\right)}. \tag{8}$$

We can then obtain a preliminary estimation by adding up the congestion measures due to different nets:

$$P(x, y) = \sum_{\text{all } k} P_k(x, y). \tag{9}$$

7.2 Detailed Estimation

In our approach, we will predict the congestion measures by using a diagonal (orthogonal to the source-sink connection) based model during the detailed estimation step. We first assume that all the nets are routed in their shortest Manhattan distances. The tiles inside the smallest bounding box of net k can be divided into $DT_k - 1$ divisions where DT_k is the shortest Manhattan distance between the source and the sink. An example is shown in Figure 4.

In this example, the tiles are divided into three divisions D_1, D_2, and D_3. Intuitively, if the net is restricted to be routed within the bounding box, the net will pass through exactly one tile in each division. We assume that the net will pass through the tiles in the same division with probabilities weighted according to $W(x, y)$ where $W(x, y)$ is computed by the following equations according to the $P(x, y)$ obtained in the preliminary estimation step.

$$W(x, y) = \begin{cases} 1 & : P(x, y) < \left(c^h_{max} + c^v_{max}\right) \\ \frac{c^h_{max} + c^v_{max}}{P(x, y)} & : \text{otherwise.} \end{cases} \tag{10}$$

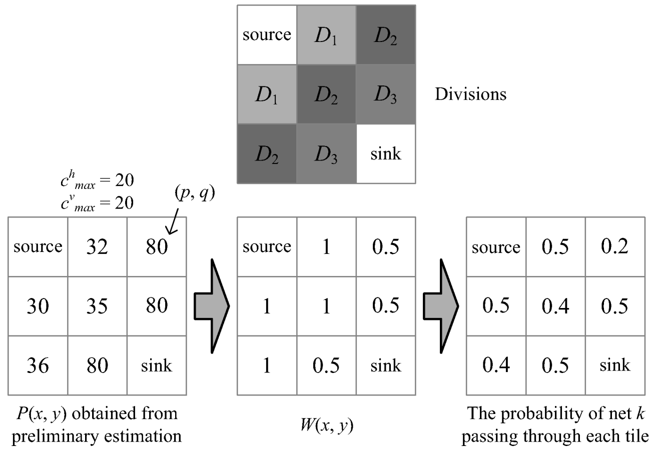

Fig. 4. An example of computing the congestion measures for a two-pin net in the detailed estimation step.

If $P(x, y)$ is smaller than the sum of c^h_{max} and c^v_{max}, the tile (x, y) is unlikely to be over-congested and so $W(x, y)$ is 1. If $P(x, y)$ is larger than the sum of c^h_{max} and c^v_{max}, that tile should have a smaller $W(x, y)$ when $P(x, y)$ is larger. It reflects the case in the routing stage that the nets will be routed to pass through less congested tiles. Hence, the probability of net k passing through (x, y), $E_k(x, y)$, can be calculated according to the weight of each tile, $W(x, y)$, by the following equation:

$$E_k(x, y) = \frac{W(x, y)}{\sum_{(i,j)\in T_k(d_k(x,y))} W(i, j)}. \tag{11}$$

In the example of Figure 4, c^h_{max} and c^v_{max} are 20. When we focus on division D_2, (p, q) the tile at the upper right corner is a congested tile according to the preliminary estimation step because $P(p, q)$ is 80, which is larger than the sum of c^h_{max} and c^v_{max}. Thus, $W(p, q)$ should be smaller than 1 and it is computed as 0.5 according to Equation (10). Hence, the probability of net k passing through (p, q), $E_k(p, q)$, is 0.2. It is smaller than the others in the same division because the tile (p, q) is likely to be overcongested. In addition, a net may pass through a tile either horizontally or vertically. We will calculate $E^h_k(x, y)$ and $E^v_k(x, y)$ by Equation (2a) and (2b) respectively.

7.3 Congestion Redistribution

In real routing, if some tiles are over-congested or some nets cannot be routed, rip-up and reroute will be performed. In our approach, we perform congestion redistribution to achieve the same purpose of moving wires from overcongested tiles to less congested tiles. We will only move around those congestion measures within the same diagonal (division). In this case, we may simulate the routing process when less congested path is always selected and the total congestion of one division can be maintained at one. An example is shown in Figure 5. In $T_k(1)$ of net k, the tile with congestion estimation 7.2 is the most congested in

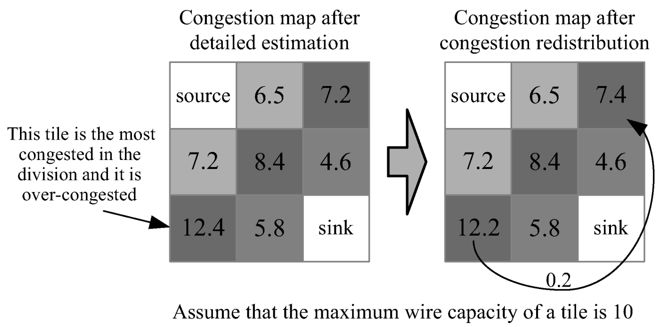

Fig. 5. An example of congestion redistribution.

this division but it is not overcongested (less than the maximum wire capacity of a tile). Thus, no action will be taken. In $T_k(2)$, the tile with 12.4 is the most congested in this division and is overcongested. Thus, we will move 0.2 (net k's contribution to the congestion measure of this tile) from this tile to the least congested tile of the same division.

In general, we will find the tile, (x_m, y_m), with the maximum vertical (horizontal) congestion and the tile, (x_l, y_l), with the minimum vertical (horizontal) congestion from each division of all the nets. If the tile with the maximum vertical (horizontal) congestion is overcongested, we will move $E_k^v(x_m, y_m)$ ($E_k^h(x_m, y_m)$) from (x_m, y_m) to (x_l, y_l). After redistribution, the summation of $E_k^v(x, y)$ ($E_k^h(x, y)$) in the same division still equals one. Thus, the assumption that each net will pass through exactly one tile in each division within the bounding box still holds.

8. BLOCKAGES

Blockages are regions with reduced routing resources. There are two types of routing blockages: partial or complete. Partial blockages block a certain number of layers, but there are still limited routing resources available. Complete blockages block all the layers, and no net can pass through those blockages. The degree of blocking, $B(x, y)$, at a tile (x, y) can be calculated by the following equation:

$$B(x, y) = \frac{\text{No. of blocked layers}}{\text{Total number of layers}}. \quad (12)$$

Then the weight of a tile (x, y), $W(x, y)$, will be updated by the following equation:

$$W(x, y) = W(x, y) \times (1 - B(x, y)). \quad (13)$$

Note that $B(x, y)$ equals one when all the layers are blocked (complete blockage).

9. EXPERIMENTAL RESULTS

In the experiments, the test cases used are the ISPD-02 suite circuits [ISPD 2002]. The length of a tile, t_l, is $40\mu m$. The detailed information of the testing

Table VI. Information of the Test Cases

Test Cases	No. of Cells	No. of Nets	No. of 2-pin Nets	No. of Tiles
*ibm*01	12506	14111	36455	57 × 57
*ibm*02	19342	19584	61615	82 × 82
*ibm*03	22853	27401	66172	89 × 87
*ibm*04	27220	31970	73889	85 × 86
*ibm*05	28146	28446	97862	60 × 60
*ibm*06	32332	34826	93366	81 × 82
*ibm*07	45639	48117	127522	97 × 97
*ibm*08	51023	50513	154377	104 × 103
*ibm*09	53110	60902	161186	118 × 118
*ibm*10	68685	75196	222371	194 × 189
*ibm*11	70152	81454	199332	130 × 129
*ibm*12	70439	77240	240520	171 × 171
*ibm*13	83709	99666	257409	141 × 141
*ibm*14	147088	152772	394044	151 × 151
*ibm*15	161187	186606	529215	170 × 169
*ibm*16	182980	190048	588775	204 × 203
*ibm*17	184752	189581	670455	182 × 182
*ibm*18	210341	201920	617777	163 × 163

circuits are shown in Table VI. Each circuit is first placed using a wirelength driven placer, Capo [Caldwell et al. 2000]. Four placement solutions are obtained for each test case. Global routing is then performed on each placement solution by a maze routing based global router [Kastner et al. 2002]. During global routing, we set wiring capacity value to simulate two environments: more congested and less congested cases. For the data sets shown in Table VII, there are about $0\% - 2\%$ overcongested tiles after global routing. Different congestion models are then used to estimate the congestion of the placed circuits and their estimations are then compared with the actual congestion measures obtained from the global router.

We compare our SMD model, Detour model and the 3-step approach with the models from Lou [Lou et al. 2001] and Westra [Westra et al. 2004]. We have implemented all the congestion models and compared the estimations with the results of the maze router. All programs were written in the C language and run on a machine (Sun Blade 1000) with $750MHz$ processor and $2GB$ memory. We will compare the congestion models by calculating the mean of error, μ and the standard deviation of error, μ_{std} according to the following equations:

$$\mu_h = \frac{\sum_{(x,y)\in T} \frac{|A^h(x,y)-E^h(x,y)|}{c^h_{max}}}{|T|}$$

$$\mu_v = \frac{\sum_{(x,y)\in T} \frac{|A^v(x,y)-E^v(x,y)|}{c^v_{max}}}{|T|}$$

$$\mu = \frac{\mu_h + \mu_v}{2} \tag{14}$$

$$\mu_{std} = \sqrt{\frac{\sum_{(x,y)\in T}\left(\left(\frac{|A^v(x,y)-E^v(x,y)|}{c^v_{max}} - \mu\right)^2 + \left(\frac{|A^h(x,y)-E^h(x,y)|}{c^h_{max}} - \mu\right)^2\right)}{|T|}}, \tag{15}$$

Table VII. Comparison on the Mean and Standard Deviation of Error of the Congestion Models for More Congested Circuits

Test Cases	c^h_{max} and c^v_{max}	Lou's Model		Westra's Model		SMD Model		Detour Model		3-step Approach	
		μ	μ_{std}	μ	μ_{std}	μ	μ_{std}	μ	μ_{std}	μ	μ_{std}
*ibm*01	27	17.36	10.01	15.06	10.05	14.32	9.57	14.05	9.58	12.63	9.45
*ibm*02	46	16.96	11.27	14.12	11.32	11.63	10.88	11.20	10.92	10.15	10.64
*ibm*03	45	32.34	10.78	27.59	10.75	24.29	10.82	22.32	11.05	20.66	10.78
*ibm*04	140	4.51	11.69	4.37	11.64	4.12	11.02	4.12	10.67	4.10	10.57
*ibm*05	70	27.21	11.23	23.27	11.37	18.28	11.19	18.10	11.54	14.22	10.67
*ibm*06	47	23.43	12.32	21.20	12.15	17.61	12.35	17.05	12.02	14.65	12.64
*ibm*07	53	13.51	12.58	12.05	12.43	10.44	11.65	10.33	11.98	9.44	12.21
*ibm*08	400	2.17	11.06	2.07	11.02	1.86	10.43	1.85	10.54	1.85	9.88
*ibm*09	50	15.76	10.78	13.62	10.68	11.43	10.97	11.32	11.43	10.34	12.08
*ibm*10	50	18.16	12.63	11.71	12.73	9.13	12.22	9.02	12.05	8.36	11.54
*ibm*11	55	14.38	12.43	11.55	12.62	10.28	11.97	10.01	11.54	9.43	12.64
*ibm*12	60	20.79	11.57	14.86	11.67	11.42	11.64	11.12	11.95	10.16	12.67
*ibm*13	60	14.67	12.12	11.97	12.69	10.21	11.00	9.98	11.34	9.35	10.68
*ibm*14	65	13.38	11.67	10.52	11.09	9.88	10.97	9.57	11.68	9.32	11.02
*ibm*15	80	17.19	10.52	12.06	10.67	10.33	11.02	10.10	10.57	9.28	10.64
*ibm*16	60	19.54	12.98	14.47	12.84	11.58	12.54	11.26	12.06	10.58	12.03
*ibm*17	80	18.56	10.67	12.70	10.67	10.57	9.87	10.23	9.77	9.69	9.44
*ibm*18	70	17.20	12.04	14.56	11.97	12.76	12.54	12.35	12.44	11.65	12.22
Average		17.06	11.58	13.76	11.58	11.67	11.26	11.33	11.29	10.33	11.21
w.r.t. Lou's		1.00	1.00	0.81	1.00	0.68	0.97	0.66	0.97	0.61	0.97

where T is the set of all tiles that either their actual congestion measures or estimated congestion measures are nonzero.

The experimental results are shown in Table VII. The values are the averages of the four placement solutions for each test case. We can see that SMD model can give smaller means in most cases than Lou's [Lou et al. 2001] and Westra's [Westra et al. 2004] models. Detour model also have smaller means and standard deviations of error than SMD model but the improvement is not significant. The accuracies can be further improved when we can simulate the rip-up and reroute operations by performing the preliminary estimation and congestion redistribution steps by the 3-step approach. For the standard deviation of the error, the results of all the models are similar.

In Figure 6, the congestion maps obtained by different congestion models and the actual one (obtained by global routing) are shown. We can see that there are many regions that are predicted as overcongested in Lou's and Westra's models and there are also a lot of empty regions in their models. However, the nets can be ripped up and rerouted to avoid passing through the overcongested regions. There is thus no overcongested region after global routing and most of the tiles in the placed region are used by some nets. In our modeling, we applied the preliminary estimation and congestion redistribution steps, and a similar congestion map can be obtained. Clearer comparisons can be illustrated by the error distributions of different congestion models in Figure 7. We can see that differences occur in the surroundings of the overcongested tiles. It is because the global routing step will rip up the nets from the overcongested tiles and reroute

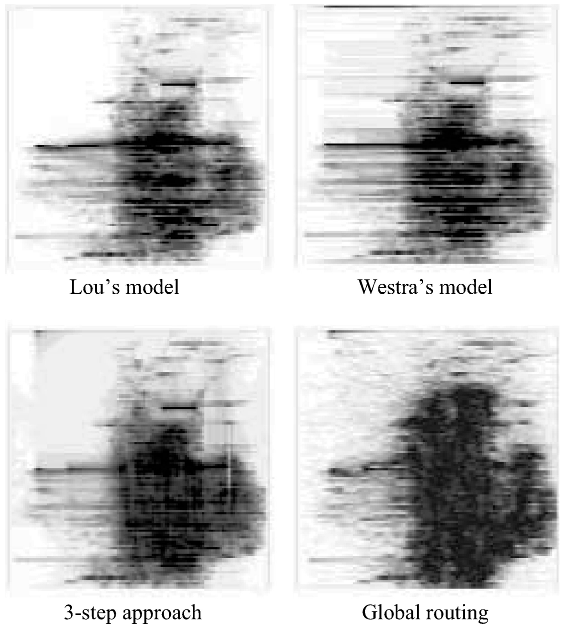

Fig. 6. Congestion maps of horizontal wires (case: ibm03).

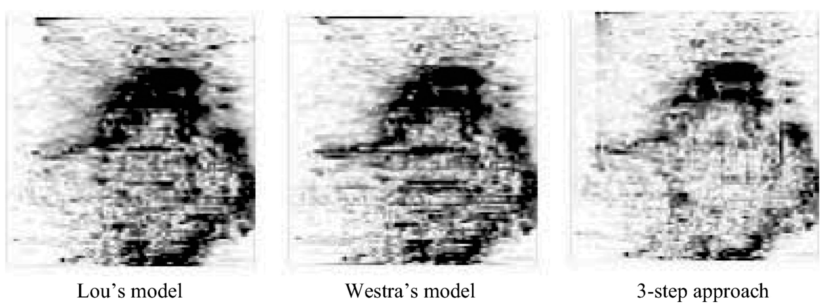

Fig. 7. Error distribution of horizontal wires (case: ibm03).

them in the less congested tiles in the surroundings. Results show that we can improve the congestion estimation accuracy in different parts of the circuit.

In addition, we have compared the runtime of different congestion models. The results are shown in Table VIII. If we apply the detailed estimation step only, the runtime is faster than both Lou's [Lou et al. 2001] and Westra's [Westra et al. 2004] models. If we also apply the preliminary estimation and congestion redistribution steps, the runtime is slower. However, it is still acceptable because a more accurate congestion model can help us to spend less time in the later routing stage.

We compare the mean of error of the congestion models with ISPD-07 suite circuits[1] and the circuits are global routed by two different global routers

[1]http:// www.sigda.org/ispd 2007/rcontest/.

Table VIII. Comparison of the Runtime of the Congestion Models

Test Cases	Lou's Model (s)	Westra's Model (s)	SMD Model (s)	Detour Model (s)	3-step Approach (s)	Global Routing [Kastner et al.] (s)
*ibm*01	0.24	0.14	0.13	0.19	0.31	190
*ibm*02	0.46	0.28	0.27	0.42	0.60	454
*ibm*03	0.92	0.58	0.53	0.92	1.10	987
*ibm*04	0.94	0.54	0.50	0.93	1.00	806
*ibm*05	1.03	0.60	0.55	0.94	1.20	1058
*ibm*06	0.64	0.35	0.34	0.59	0.77	642
*ibm*07	1.16	0.87	0.67	1.08	1.44	1206
*ibm*08	1.46	1.00	0.94	1.53	1.96	2021
*ibm*09	1.50	1.02	0.95	1.43	2.02	2217
*ibm*10	5.09	4.56	4.20	6.23	7.93	8820
*ibm*11	2.25	1.61	1.51	2.49	3.07	3021
*ibm*12	6.00	5.49	5.08	8.03	9.60	10543
*ibm*13	2.93	2.05	1.90	2.98	3.91	4680
*ibm*14	4.45	3.14	2.94	4.87	5.93	9480
*ibm*15	6.99	5.51	5.11	7.61	10.21	14220
*ibm*16	8.01	6.32	5.90	9.39	11.68	15684
*ibm*17	9.77	7.81	7.27	11.36	14.24	20547
*ibm*18	4.84	3.14	2.92	4.95	6.43	11235
Ave.	3.26	2.50	2.32	3.67	4.63	5990

Table IX. Comparison on the Mean of Error of the Congestion Models When the Circuit is Global Routed by AMGR

Test Cases	Lou's Model (μ)	Westra's Model (μ)	SMD Model (μ)	Detour Model (μ)	3-step Approach (μ)
*adaptec*1	22.55	21.28	20.78	21.77	19.88
*adaptec*2	18.77	22.09	22.50	23.26	22.37
*adaptec*3	8.98	5.54	4.74	5.02	4.23
*adaptec*4	6.82	3.79	3.49	3.54	3.31
*adaptec*5	21.06	13.64	12.07	12.94	9.68
*newblue*1	6.60	5.42	4.22	4.37	4.03
*newblue*2	5.53	3.75	3.08	3.09	3.02
Ave. (w.r.t. Lou's)	1.00	0.78	0.70	0.73	0.66

Table X. Comparison on the Mean of Error of the Congestion Models When the Circuit is Global Routed by MaizeRouter

Test Cases	Lou's Model (μ)	Westra's Model (μ)	SMD Model (μ)	Detour Model (μ)	3-step Approach (μ)
*adaptec*1	22.50	20.95	20.40	21.40	19.51
*adaptec*2	18.92	22.12	22.51	23.26	22.37
*adaptec*3	8.72	5.43	4.68	5.06	4.18
*adaptec*4	6.71	3.80	3.31	3.41	3.13
*adaptec*5	23.83	16.30	14.68	15.60	12.16
*newblue*1	6.67	5.16	4.21	4.32	4.03
*newblue*2	5.29	3.64	2.99	3.01	2.94
Ave. (w.r.t. Lou's)	1.00	0.78	0.71	0.73	0.67

(AMGR [Xiao et al. 2008] and MaizeRouter [Moffitt 2008]). The results are shown in the Table IX and Table X. We can see that the 3-step approach also have smallest mean of error among different congestion models. Besides the circuit *adaptec*2, the results are very consistent. Although these two global routers apply two different approaches (AMGR mainly applied maze routing and MaizeRouter mainly applied Steiner tree routing and edge shifting), we can have similar results.

10. CONCLUSION

To conclude, we have studied and developed three congestion models to estimate congestion. Results show that the estimation results of our approaches are always more accurate than the previous congestion models. The SMD model is the fastest estimation method but it gives less accurate results. The Detour model is developed to consider detoured routing but the improvement on the estimation accuracy is found to be non-significant. The 3-step approach is developed and implemented to simulate the global routing, detailed routing and rip-up and reroute process in the real routing step and it makes significant improvement on the prediction accuracy efficiently.

REFERENCES

CALDWELL, A. E., KAHNG, A. B., AND MARKOV, I. L. 2000. Can recursive bisection produce routable placements? In *Proceedings of the 37th ACM/IEEE Design Automation Conference*. 477–482.

CHANG, C. C., CONG, J., PAN, D. Z., AND YUAN, X. 2000. Interconnect-driven floorplanning with fast global wiring planning and optimization. In *Proceedings of the SRC Technical Conference*.

CHEN, H. M., ZHOU, H., YOUNG, F. Y., WONG, D., YANG, H. H., AND SHERWANI, N. 1999. Integrated floorplanning and interconnect planning. In *Proceedings of the IEEE Internation Cnference on Computer-Aided Design*. 354–357.

ISPD. 2002. http://vlsicad.eecs.umich.edu/bk/ispd02bench/.

ISPD. 2008. http://www.sigda.org/ispd2007/rcontest/.

KAHNG, A. B. AND XU, X. 2003. Accurate pseudo-constructive wirelength and congestion estimation. In *Proceedings of the International Workshop on System-level Interconnect Prediction*. 61–68.

KASTNER, R., BOZORGZADEH, E., AND SARRAFZADEH, M. 2002. Pattern routing: Use and theory for increasing predictability and avoiding coupling. In *IEEE Trans. Comput.-Aid. Des. Integr. Circ. Syst.*

KUSNADI AND CAROTHERS, J. D. 1999. A method of measuring nets routability for MCM's general area routing problems. In *Proceedings of the International Symposium on Physical Design*. 186–194.

LAI, S. T. W., YOUNG, E. F. Y., AND CHU, C. C. N. 2003. A new and efficient congestion evaluation model in floorplanning: Wire density control with twin binary trees. In *Proceedings of the Design, Automation and Test in Europe Conference and Exhibition*.

LOU, J., KRISHNAMOORTHY, S., AND SHENG, H. S. 2001. Estimating routing congestion using probabilistic analysis. In *Proceedings of the Internation Symposium on Physical Design*. 112–117.

MA, Y. C., HONG, X. L., DONG, S. Q., CHEN, S., CAI, Y. C., CHENG, C. K., AND GU, J. 2003. Dynamic global buffer planning optimization based on detail nlock locating and congestion analysis. In *Proceedings of the ACM/IEEE Design Automation Conference*. 806–811.

MOFFITT, M. D. 2008. Maizerouter: Engineering an effective global router. In *Proceedings of the ASP-ACM/IEEE Design Automation Conference*. 226–231.

PAN, M. AND CHU, C. 2006. Fastroute: a step to integrate global routing into placement. In *Proceedings of the IEEE Internation Cnference on Computer-Aided Design*. 464–471.

Saeedi, M., Zamani, M. S., and Jahanian, A. 2006. Prediction and reduction of routing congestion. In *Proceedings of the International Workshop on System Level Interconnect Prediction*. 72–77.

Sham, C. W. and Young, E. F. Y. 2005a. Congestion prediction in early stages. In *Proceedings of the International Workshop on System Level Interconnect Prediction*. 91–98.

Sham, C. W. and Young, E. F. Y. 2005b. Congestion prediction in floorplanning. In *Proceedings of the ASP-ACM/IEEE Design Automation Conference*. 1107–1110.

Taghavi, T., Dabiri, F., Nahapetian, A., and Sarrafzadeh, M. 2007. Tutorial on congestion prediction. In *Proceedings of the International Workshop on System Level Interconnect Prediction*. 15–24.

Wang, M. and Sarrafzadeh, M. 2000. Modeling and minimization of routing congestion. In *Proceedings of the ASP-ACM/IEEE Design Automation Conference*. 185–190.

Wang, M., Yang, X., and Sarrafzadeh, M. 2000. Congestion minimization during placement. In *IEEE Trans. Comput.-Aid. Des. Integr. Circ. Syst.* 1140–1148.

Westra, J., Bartels, C., and Groeneveld, P. 2004. Probabilistic congestion prediction. In *Proceedings of Internation Symposium on Physical Design*. 204–209.

Xiao, L., Li, L., and Qian, Z. 2008. *International Society for Physical Design Routing Contest (ISPD'08)*.

Yang, X. J., Kastner, R., and Sarrafzadeh, M. 2002. Congestion estimation during top-down placement. In *IEEE Trans. Comput.-Aid. Des. Integr. Circ. Syst.*, 72–80.

Received March 2007; revised November 2007, May 2008; accepted July 2008