LIMITATIONS AND OPPORTUNITIES FOR WIRE LENGTH

PREDICTION IN GIGASCALE INTEGRATION

A Dissertation
Presented to
The Academic Faculty

by

Pranav Anbalagan

In Partial Fulfillment
of the Requirements for the Degree
PhD in the
School of Electrical and Computer Engineering

Georgia Institute of Technology
May 2007

LIMITATIONS AND OPPORTUNITIES FOR WIRELENGTH

PREDICTIOIN IN GIGASCALE INTEGRATION

Approved by:

Dr. Jeff Davis, Advisor
School of Electircal & Computer
Engineering
Georgia Institute of Technology

Dr. Sung Kyu Lim
School of Electrical & Computer
Engineering
Georgia Institute of Technology

Dr. James Meindl
School of Electrical & Computer
Engineering
Georgia Institute of Technology

Dr. Paul Kohl
School of Chemical & Biomolecular
Engineering
Georgia Institute of Technology

Dr. Scott Wills
School of Electrical & Computer
Engineering
Georgia Institute of Technology

Date Approved: December 7, 2006

To my grandfathers

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

Wires have become a major source of bottleneck in current VLSI designs, and wire length prediction is therefore essential to overcome these bottlenecks. Wire length prediction is broadly classified into two types: macroscopic prediction, which is the prediction of wire length distribution, and microscopic prediction, which is the prediction of individual wires. The objective of this thesis is to develop a clear understanding of limitations to both macroscopic and microscopic a priori post-placement pre-routing wire length predictions, and thereby develop better wire length prediction models.

Investigations carried out to understand the limitations to macroscopic prediction reveal that, in a given design (i) the variability of the wire length distribution increases with length and (ii) the use of Rent's rule with a *constant Rent's exponent p*, to calculate the terminal count of a given block size, limits the accuracy of the results from a macroscopic model. Therefore, a new model for the parameter p is developed to more accurately reflect the terminal count of a given block size in placement, and using this, a new more accurate macroscopic model is developed. In addition, a model to predict the variability is also incorporated into the macroscopic model.

Studies to understand limitations to microscopic prediction reveal that (i) only a fraction of the wires in a given design are predictable, and these are mostly from *shorter nets with smaller degrees* and (ii) the current microscopic prediction models are built based on the assumption that a single metric could be used to accurately predict the individual length of all the wires in a design. In this thesis, an alternative microscopic model is developed for the predicting the shorter wires based on a hypothesis that there

are multiple metrics that influence the length of the wires. Three different metrics are developed and fitted into a heuristic classification tree framework to provide a unified and more accurate microscopic model.

CHAPTER 1

INTRODUCTION

The size of the worldwide electronics industry is estimated to be at roughly $1.3 trillion which has grown at roughly 7.1% for the last two decades. Practically every product of this industry is built out of the $300 billion semiconductor industry, which grew at 24.8% in 2004 and has been growing at roughly 15.1% during the past two decades. And nearly 28% of the $300 billion semiconductor industry is the Application Specific Integrated Circuit (ASIC) Industry. To put the economic significance of the size of these numbers in perspective the total world Gross Domestic Product (GDP) value is estimated at roughly $46 trillion, just 39x times the size of the electronic industry, and the world GDP has grown at just 3.2% for the past two decades. However, in order to maintain this growth, one must challenge the frontiers of semiconductor chip design.

With the advancements in semiconductor technology however, a bottleneck has been developing in the design process due to the influence of the wires in the overall specifications of a design [1]. One of the primary causes of this bottleneck is the lack of a clear perspective of the influence of the wires on the design's specifications during the early stages of the design process. If the length of the wires could be predicted early on during the design process, it could help overcome some of the major bottlenecks. Therefore, in this research an attempt has been made to identify the limitations to wire length prediction, and based on the knowledge gained through the study of limitations, new wire length prediction models are developed.

The core of the semiconductor design process is the Very Large Scale Integrated (VLSI) chip design flow, and the wire length problem manifests itself during this process. Therefore, in the first section of this chapter an overview of the VLSI Design flow is provided, followed by the subsections detailing the influence of the wires on the various

design specifications and the motivation for wire length prediction. In order to predict the length of the wires, it is essential to model the length of the wires, and subsequently the second section of the chapter defines the terminology used in various wire models. The third section then provides an overview of the history of various approaches to the wire length prediction problem. Finally, the chapter is concluded with an outline of the approach used in this research along with the assumptions and the basic experimental setups used in this work.

VLSI Design Flow

The design of a VLSI chip starts with setting target specifications for the chip and ends with a packaged chip after several design stages. Each design stage is aimed at simplifying the complexity of the design process to a manageable level. The various stages involved in this process are explained in this section along with a simplified design flow depicting the stages in Figure 1.1.

1. System specification: Setting the target specifications of the design is the first step of the design process. The specifications include identification of the overall functionality of the chip, its performance requirements and its target physical dimensions. These are derived based on market requirements and economic feasibility. Performance specifications will usually include the specifications of speed, power consumption, heat dissipation, etc… The physical and performance specifications are set based on the limits of the available technology.

2. Architectural Design: The second major step in the design process is to decide on the architectural specification of the chip. This includes decisions such as the instruction set, the number of pipeline stages, the number of ALU and or other such functional units and how they should be connected to each other.

Various architectural possibilities are explored at this stage to choose the best architecture to meet the target specifications.

3. Functional Design: Functional design involves selection of the various functional units and their interrelationships in terms of timing, power, heat dissipation, noise etc… to meet the target specifications. For example, this step involves decisions about the inputs to a particular functional unit, the outputs of the unit, the connections to other functional units, the sequence in which the various functional units have to operate in order for the design to meet the targeted behavior of the design.

4. Logic Design: The functional blocks developed at the previous stage are decomposed into Boolean expressions and is represented using the Register Transfer Level (RTL) description. This entails making decisions about the bit-by-bit logical details of the functional blocks and the corresponding control signal details. The Boolean expressions are then optimized, using a process called synthesis, to yield the smallest logic design for the given behavioral design. The result is a gate level description of the design.

5. Circuit Design: The circuit design stage involves designing circuits for the gates used in the design while considering the requirements of functionality, performance and physical specifications of the design. This stage produces transistor level details from the gate level description created during the logic design stage. This stage brings into view the detailed physics of the transistors and hence the performance of the gates in the design through the circuit level simulation.

6. Physical Design: Physical design is the process of physically allocating space for transistors and wires in a multilevel layout in which the given design will be manufactured. Due to the complexity of this stage, it is further subdivided into various steps viz., floorplanning, partitioning, placement and routing.

Only at the end of these stages does the designer have a clear view of the layout of wires present in the design and therefore their dimensional attributes.

a. Floorplanning is the process of planning where each of the major architectural and or functional blocks or partitions should be placed in the layout.

b. Partitioning is the process of dividing the circuit into several groupings (i.e. partitions) to bring the cells that are highly connected closer together by including them within the same partition.

c. Placement is the process of assigning an optimal location for each cell or gate in a design block or partition to the area allocated for that particular design block or partition in the layout.

d. Routing is the process of assigning an optimal location for each wire in the multiple levels of metal available in the layout.

7. Fabrication: The layout details obtained from the previous stage is first transformed into a set of photolithographic masks. These masks are then used to guide the diffusion, deposition or removal of various chemicals in the substrate to build the multiple layers of chemicals which together form an electronic chip. Several hundred chips are manufactured on a single wafer substrate, which is then diced into individual chips. Only during testing at the end of this stage does the fabrication induced effects on chip performance come to light for the first time.

8. Packaging: The individual chips are then packaged using the appropriate packaging technology or used directly in Multi-Chip Modules. The final chip is then tested against the target specifications.

```
┌─────────────────────────────┐
│     System Specification     │
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│     Architectural Design     │
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│       Functional Design      │
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│         Logic Design         │
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│        Circuit Design        │
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│       Physical Design        │
│  ┌──────────┐                │
│  │ Floorplan│                │
│  │  ┌───────┴──┐             │
│  │  │Partition │             │
│  └──┤  ┌───────┴──┐          │
│     │  │Placement │          │
│     └──┤  ┌───────┴──┐       │
│        │  │ Routing  │       │
│        └──┤          │       │
│           └──────────┘       │
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│          Fabrication         │
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│          Packaging           │
└─────────────────────────────┘
```
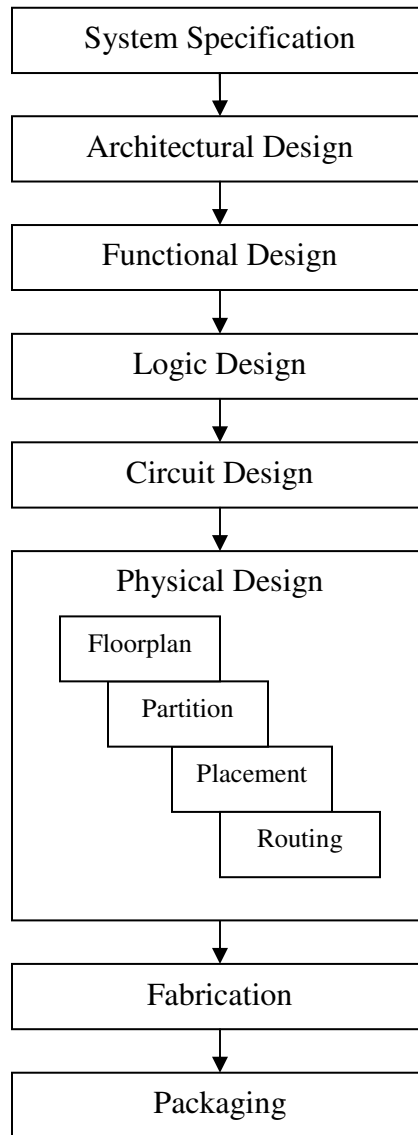
Figure 1.1 Various design phases involved in a typical VLSI design flow

The above step-by-step VLSI design process allows for a design to be evaluated at each stage, to check if the target specifications will be met at the end of the design process. In a case where the evaluation at the end of a design stage reveals that the targeted specifications cannot be met, additional design iteration is performed starting from an earlier stage after making some modifications to the design. However, increasing the number of iterations increases the cost of the design and its time-to-market, resulting in reduced marketability of the product.

Impact of Scaling

As the technology advances over time, the functionality of the chip is increasing with a trend that follows Moore's law [2], according to which the number of transistors in a chip doubles every 18 months. This rapid increase in the number of components within a chip is accompanied by a decrease in the feature size of the components and wires within the chip. Further, the connected gates are spread relatively farther apart in the layout as a consequence of the increased number of gates in the design. Consequently, as the size of the chips increase, so do the number of wires inside the chip and the relative length of the wires with respect to the size of the gates inside the chip. Coupled with the decreasing feature size, the resistance of the wires starts to increase dramatically since its cross-sectional area decreases and the number of longer wires increases from generation-to-generation. Further, the capacitance contribution of the wires also increases with increasing length. Meanwhile, the reduction in the feature sizes causes the transistors to switch faster. The combined interaction of these changes has allowed the wires to strongly influence every aspect of a design specification in the following ways:

1. Impact on Physical Specifications: The number of wires and their relative lengths increase from generation-to-generation and from design-to-design. To accommodate this, the number of metal levels allowed in the manufacturing technology also grows. But this growth rate is sometimes not commensurate with the amount of wire space required to accommodate the wires within the targeted physical dimensions of the chip based solely on gate occupied area. Furthermore, the increasing complexity of the designs increases the number of wires in some regions of the design, which in turn increases the wire congestion in those regions. This congestion, coupled with more wires getting longer, makes them more susceptible to noise. Therefore, increased spacing and shielding is needed between wires to alleviate possible noise problems. Consequently the wires are spread out within the available metal levels. As

the wires are being spread out, at some point the area of the wires start to dominate over the area of the gates. Moreover, repeaters inserted to speed up the longer wires will also start pushing the layout size [3]. The result is that the wires limit the physical specification (area) of the chip.

2. Impact on Timing/Speed Specifications: As the resistance and capacitance of the wires increase, the time it takes for a signal to be transmitted through the wires increases. For example, wire delay degradation per scaled micron is of the order of 1.4x for every generation [3]. Further, since the time it takes for the signal to be transmitted through the transistors reduces, the propagation delay through the wires are playing a dominant role in the timing of the chip, which in turn is controlled by its length. A widely employed methodology used to reduce the propagation delay through longer wires is to insert a number of repeaters on the longer wires, where the repeaters function by allowing the signals to be accelerated through the wires. The number of repeaters that can be inserted, and hence the speed-up achieved, depends upon the length of the wires.

3. Impact on Power Specifications: As the number of wires and the wire length increases from technology generation-to-generation, in order to overcome the propagation delay problem, an increasing number of repeaters are employed The number of repeaters deployed increases with the number and length of the longer wires. For example, it is predicted that 50% of the cell area at 32nm technology generation will be filled with repeaters [3]. Since each of these repeaters consume static and dynamic power, the power consumption of a design will be dominated by the number of repeaters in the design. Because the number of repeaters in the design depends upon the length of the wires in the design, the wires have started to play a major role in the power consumption specification of the chip [4].

4. Impact on Thermal Specifications: A large fraction of the energy is dissipated as heat during each switching event in a design. And as the wire resistance per micron doubles with every technology generation [3], more of this heat dissipation occurs through the wires. Further as the designs get more complex, some of the regions are highly congested leading to possible hot spots in the design. For example, in [5] it is argued that a non negligible amount of heat dissipation occurs through the wires. Consequently, the thermal specifications of the system are also altered by the wires present in the system.

5. Impact on Noise Specifications: With the increase in the length and the number of the wires, the wires are more likely to be adjacent to a number of other wires over long distances in the system. These wires are coupled with each other through a coupling capacitance and inductance and can thus induce noise on each other during switching transitions. Thus the wires are bound to influence the noise specifications of a chip design as well.

6. Impact on Reliability Specifications: With the increase in wire length and wire resistance, wires dissipate energy in the form of heat. As the wires get overheated this could lead to problems such as electromigration, which could affect the conductivity of the wires and thus the reliability of the design.

7. Impact on Functional Specification: The influence of wires as explained above in each of the performance specifications in turn influences the functionality of the overall design. For example, an unwanted time delay of the signal through the wires could change the function of clocked logic that is set up to run with a different timing specification. In addition, unwanted noise from wires could result in an undesired effect in the functionality of the design.

Thus it can be seen that the wires play a crucial role in determining the various specifications of a given design.

Motivation for Wire Length Prediction

In the process of a VLSI design, the designer needs to explore the design space for each stage of the design. For example, at the architectural design phase the designer needs to evaluate the impact of the different architectural choices on the final target specifications. A highly parallel architecture choice may result in a netlist structure that is very different from a less parallel architecture choice. One architecture might result in a fewer number of gates and lesser interconnection complexity between the gates than another. Consequently, the changes in these architectural design choices could affect the final specifications of the design. Similarly, the design choices in each of the other design stages such as alternative behavioral designs, logic designs, circuit designs, partitioning strategies, floorplanning strategies, placement and routing strategies have to be explored. Each of the design choices will have their own unique attributes that will influence the final characteristics of the design.

Following the exploration of the various design choices, a decision has to be made as to which choice to use for the continued development of the design. In order to make such a decision, it is necessary to evaluate if a design choice will meet the target specifications at the end of the design process. If the design choice meets the targeted objectives, then the design is said to converge. If the evaluation stage indicates that design convergence cannot be achieved, the designer needs to go back to the previous design stages to alter the earlier design choices.

To evaluate if a design choice will meet the target specifications of the design, it is necessary to model the influence of the wires in the specifications. This is because the wires practically influence every one of the target specification metrics. And for the most part, the magnitude of this influence is dependent primarily upon the physical characteristics of the wires because its electrical attributes are dependent upon its physical characteristics. However, the physical attributes are only available at the end of the physical design stage when the gates have finally been placed in the physical layout

and the wires have been routed through the multilayer metals. Therefore, it is impossible to include the exact effect of wires on the various specification metrics at the earlier stages of the design. At the same time any evaluation of the specification metrics at the earlier stages of the design flow without an appropriate model to include the impact of wires renders them useless. This lack of information in the design specifications at the earlier stages of the design flow is one of the major bottlenecks to design convergence leading to iterations in the design flow, as shown in Figure 1.2. A prediction model of the wire attributes could, therefore, allow the designer to design by including the effect of wires from the very early stages of the design flow, and thereby could help achieve a much faster design convergence.
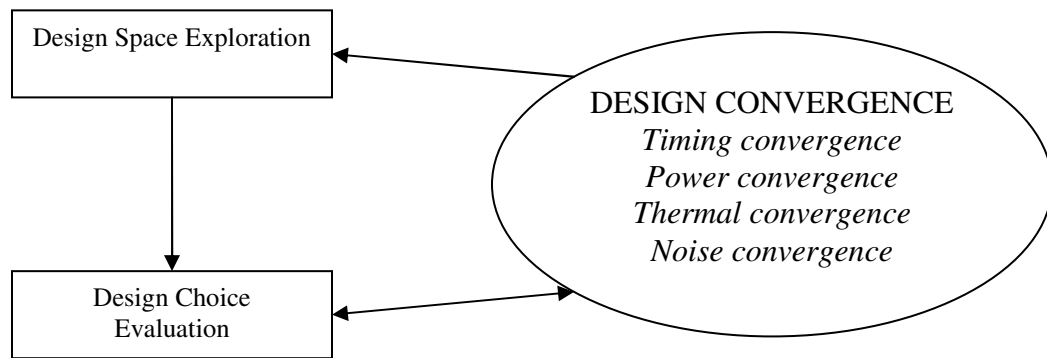
Figure 1.2 Relation between design space exploration and design choice evaluation

Further, the advancements in chip design and manufacturing technology will create newer design challenges. To understand and predict these newer challenges, it is necessary to develop technology extrapolation models. Such early extrapolations of future technological problems will enable the designers to be better prepared to design future systems. Due to the increasing impact of wires with scaling, it is even more important to include models that would incorporate the effect of wires on the extrapolated design problems. Since the impact of wires on a design is strongly linked to wire length, it is necessary to develop wire length prediction models.

Wire Model Terminologies

A *wire* is an equipotential connection that conducts a common electrical signal from one or more source (driver) gates to the one or more of its sink (driven) gates. During the chip design process, the design is often represented in the form of a graph with the gates or functional blocks being represented as the vertices of the graph and the edges or hyperedges between the vertices representing the wires connecting the corresponding blocks or gates. These graphical models of the wire can be classified into a *net* model, *interconnect* model and *connection* model.

A *net* is the multi-terminal hyperedge model of the wire that connects all the gate nodes representing the source and sink gates of the common signal. Since the source of the common signal could be from more than one gate and since it could drive more than one gate, a net model of a given wire may have more than two terminals. This is the most accurate representation of the wire in a graph. For example, consider a signal wire with gate G1 as its source and gates G2, G3 and G4 as its sink. Such a wire can be represented as a multi-terminal hyperedge between the four gates as shown in the Figure 1.3(a). The first estimate of the length of the wire is possible only after the gates are placed on the layout. Now if the gates G1, G2, G3 and G4 are placed in a layout that is in the form of a grid, as shown in Figure 1.3(b), then the post-placement pre-routed length of the net can be computed using any of the following length estimation models shown in Figure 1.3(c), 1.3(d) and 1.3(e). Figure 1.3(c) represents the Rectilinear Steiner Tree based representation of the net. The length of the net based on the rectilinear Steiner tree is the sum of the length of the horizontal and vertical edge segments of the rectilinear Steiner tree. Figure 1.3(d) shows the spanning tree based representation of a net. The length of the net based on a spanning tree is the sum of the manhattan lengths of the edges in the spanning tree. And finally Figure 1.3(e) shows semi-perimeter bounding box based estimation of the net length. Here the length of the net is estimated as half the perimeter of the smallest rectangular box that encloses all the nodes of a given net. While

11

an estimation based on minimum rectilinear Steiner tree gives the minimum actual possible length between the nodes of the net, it is hard to find the actual minimal cost rectilinear Steiner tree, and the actual routed wire length may be greater than this value. However, finding a minimal rectilinear spanning tree is easier and is a close approximation to minimal rectilinear Steiner tree, but it still takes some computational power to estimate them. On the other hand, estimation of the semi-perimeter bounding box length takes significantly less computational power; but the pitfall of this method is that it could be inaccurate for nets with a large number of terminals.

(a) Net model of a wire   (b) Layout representation

(c) Steiner tree   (d) Spanning tree   (e) Semi-perimeter
                                                     bounding box
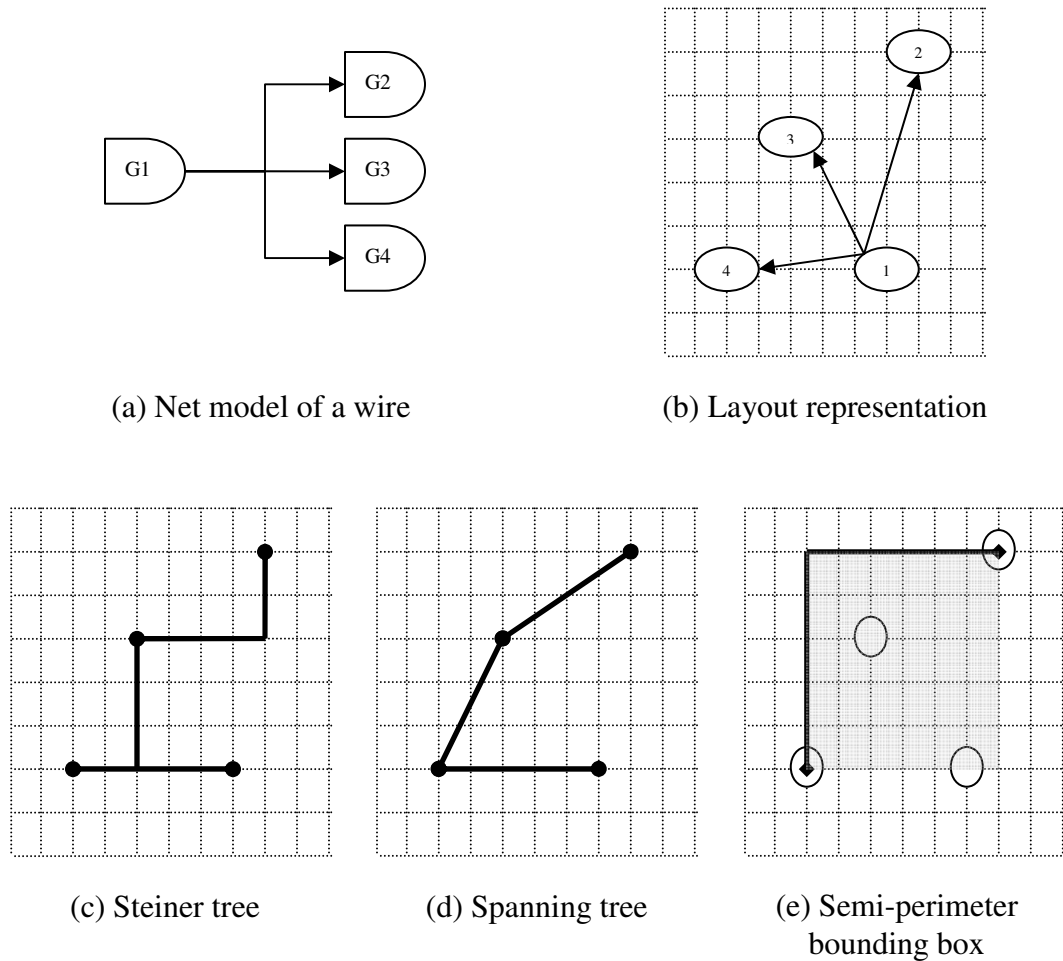
Figure 1.3 Net model of a wire and corresponding length estimation models

An *interconnect* model of the wire, see Figure 1.4(a) is a much simpler representation of the wire than a net. In this model, the wires with more than two terminals are decomposed into several two terminal edges. The two terminal edges run only between each pair of a source and sink gate in the net. The length estimates are then made individually for each interconnect as the manhattan distance between the two terminals of the two terminal edges as shown in figure 1.4(c). This model is similar to a spanning tree based model. The difference lies in the fact that the two terminal edges between the gates are treated individually, and also this model eliminates the computation required to find the minimal spanning tree by limiting the edges between source and sink gates of a net.
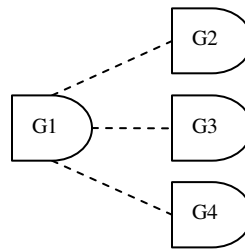
(a)
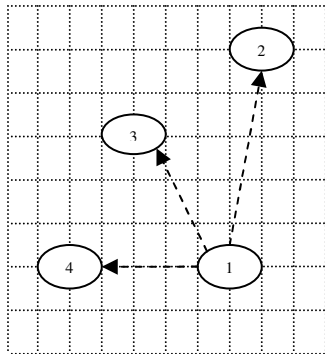
(b)
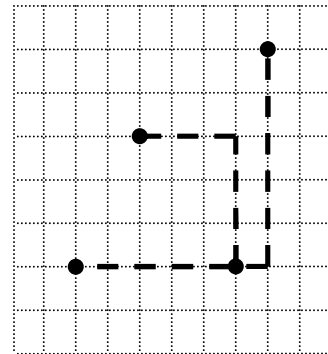
(c)

Figure 1.4 Interconnect model of a wire and corresponding length estimation model

Finally, a *connection* model of a wire is shown in Figure 1.5(a). It is a simple representation similar to interconnect model in the sense that the hyperedge is decomposed into several two terminal edges. But this model includes a two terminal edge in between every pair of terminal gates in the net. This is unlike in the interconnect model, where the two terminal edge runs solely between a source-sink gate pairs. The length of each connection is estimated separately as the manhattan distance in between the two terminal nodes of the connection placed as shown in figure 1.5(b). Based on the application requirements, all of the above models have been used in the literature to estimate wire lengths.

(a)   (b)
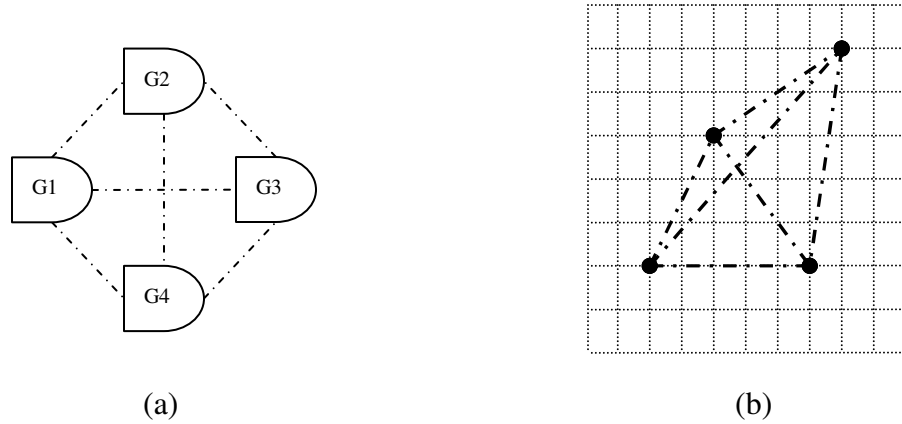
Figure 1.5 Connection model of a wire and corresponding length estimation model

History of Wire Length Prediction

Literature research reveals wire length prediction models that provide both the microscopic and macroscopic perspectives of wires in a design. *Microscopic* perspectives are provided by models that aim to predict individual wire length, and *macroscopic* perspectives are provided by models that aim to provide a more global view of the wires by predicting statistics such as the wire length distribution. However most of these methods have their limitations and are lacking in accuracy due to the difficulty involved in prediction of wire lengths.

14

Early work in interconnect length prediction was carried out in 1979 by Donath to help estimate the wiring space requirements, delay values and power dissipation [6]. In this work an upper bound for average interconnection length is estimated. It is based on hierarchically applying the relationship between the number of terminal and the number of gates for a given circuit size. The interpretation of this terminal-to-gate relationship known as Rent's rule is described in [7], [8]. According to this interpretation, the number of input and output terminals of nets T leaving a block containing C cells is given by the expression in (Eq1.1), where k and p are empirical constants known as the Rent-coefficient and the Rent- exponent respectively.

$$T = kC^p \qquad\qquad (Eq1.1)$$

While [6] gives an upper bound approximation on the average length of all interconnects in the circuit, in [9] Donath develops the early model of interconnect length distribution based on the Rent's rule. According to this model the number of interconnects of length l is given by the expression in Eq1.2, where α is a constant.

$$I[l] = \frac{\alpha C\left(1 - 4^{p-1}\right)}{2} l^{2p-3} \qquad\qquad (Eq1.2)$$

Unlike the macroscopic prediction by Donath, the method in [10], [11] increases the resolution of length prediction by predicting a different net length for each net degree. The length of a net of a given net degree is determined by considering all possible pin configurations of the net over its average neighborhood size, where the size of the neighborhood is measured as the number of gates present in the neighborhood. The neighborhood of a net refers to the cells directly connected to this net and its first-level neighbors. For example, Figure 1.6 shows a net (reference net) and its neighborhood as defined in [10], [11]. The first-level neighbors are those cells that are connected to the cells of the reference net through exactly one other net. In [12] Hamada et al., extend the neighborhood concept to include the second-level neighbors, where the second level neighbors of a net are those cells that are connected to the cells of the reference net

separated by two nets. The method in [12] predicts the connection length distribution of nets of different net degrees by assuming that the length distribution is a form of Weibull distribution and the results [12] are shown to be better than the model in [11]. However, it should be noted that the two methods use different wire length estimation models, and therefore, direct comparisons may not be accurate. In spite of the differences both the models are observed to perform poorly for the largest benchmark tested. One of the main drawbacks of these methods in [10], [11] and [12] is that they ignore the effect of those cells and nets that are outside this small neighborhood.
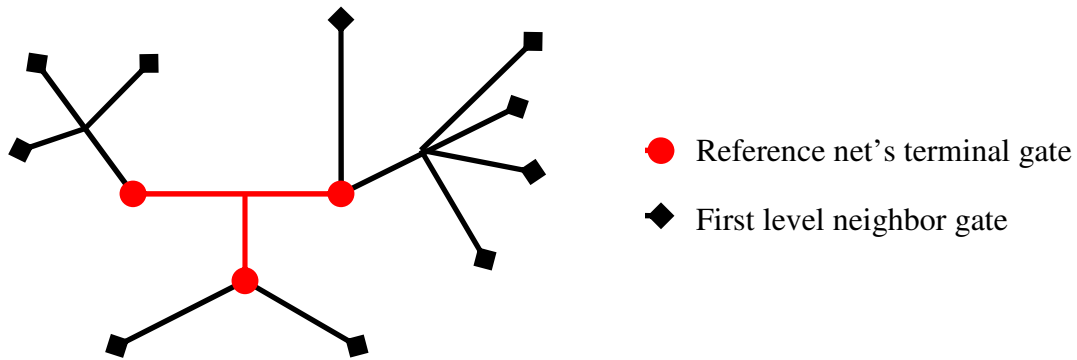
Figure 1.6 A net and its first level neighborhood

The authors in [13] make one more step forward in microscopic length prediction by estimating the mean and variance of an interconnect length for each interconnect separately. The interconnect lengths are calculated based on structural attributes extracted from its local neighborhood and some global parameters. The local attributes are the interconnect weight (*IW*) and the neighborhood population at the third-level of the neighborhood (*Angh3*). The interconnect weight is the sum of the number of unique nets connected to the terminals of an interconnect and degree of the interconnect's net. The global attributes are the total number of cells in the circuit *C* and the total number of nets in the circuit *N*. This method establishes a linear relationship between the interconnect weight and both the mean and variance of the interconnect length as shown in expressions (Eq1.3 and Eq1.4). It involves extracting data from known placement results

and then fitting them to a straight line to get the appropriate coefficients (*am*, *bm*, *av*, *bv*) for calculating the mean and variance for other circuits.

$$Mean = am + bm \cdot Angh3 \cdot IW \tag{Eq1.3}$$

$$Variance = av + \frac{bv \cdot Angh3 \cdot IW \cdot C}{N} \tag{Eq1.4}$$

The last decade has seen a number of research efforts targeting the prediction of interconnect length distribution of a complete design. In [14] a new interconnect length distribution model was developed by taking into account the number of possible sites available for an interconnect to occupy at a particular length. The interconnection length distribution for a circuit with C cells is given by the product of the interconnect occupational probability function (*Ip*[*l*]) and the site density function (*M*[*l*]) as given by the expression in (Eq1.5). In the expression for occupational probability (Eq1.7) the parameters *k* and *p* are the Rent-parameters and the parameter α gives the fraction of terminals that are sinks (Eq1.8). This parameter α is used for converting the number of net terminals between blocks into number of interconnects between blocks.

$$I[l] = M[l] \cdot Ip[l] \tag{Eq1.5}$$

$$for 1 \leq l < \sqrt{C}:$$
$$M[l] = \left(\frac{l^3}{3} - 2l^2 \sqrt{C} + \frac{1}{3}l(6C - 1) \right)$$
$$for \sqrt{C} \leq l < \left(2\sqrt{C} - 2 \right): \tag{Eq1.6}$$
$$M[l] = \left(-\frac{l^3}{3} + 2l^2 \sqrt{C} - \frac{1}{3}l(12C - 1) + \frac{2}{3}\sqrt{C}\left(2\sqrt{C} - 1\right)\left(2\sqrt{C} + 1\right) \right)$$

$$Ip[l] \cong \alpha k \frac{p}{2}(2 - 2p)l^{2p-4} \tag{Eq1.7}$$

$$\alpha = \frac{fanout_{avg}}{1 + fanout_{avg}} \tag{Eq1.8}$$

The above model considered the entire system as a single block. In [15] however a model that takes into account the hierarchy of the system was developed by Stroobandt

et al. This was done to more closely approximate the multilevel partitioning and placement process which divides the system into hierarchical blocks. In spite of this difference, the model in [15] is still similar to the model in [14] both of them calculate the wire length distribution as a product of the site density function and occupational probability function. Further, the model in [15] also uses the Rent-exponent to calculate the number of interconnects at each hierarchical level and their corresponding length distribution. The overall length distribution is then obtained by the sum of the interconnect length distributions over all the hierarchical levels. This concepts from this model was then extended in [16] into a new model, also developed by Stroobandt, to predict the net length distributions. The length predicted in this model is the Steiner tree length of the nets.

Since these distribution models were heavily dependent on the accuracy of the Rent's exponent, Rent's exponent extraction methods were studied in [17], [18] and [19]. It was observed that the Rent's exponent extracted from placement was always greater than the Rent's exponent extracted from partitioning. Further, it was also shown in [18] that among the two different Rent's exponents extracted from placement, the average local Rent's exponent was greater than the placement Rent's exponent. It was also suggested that the hierarchy-unaware wire length estimation model of [14] should use the average local Rent's exponent for interconnect length distribution calculation, and the hierarchy-aware wire length estimation model in [15] should use the partitioning Rent's exponent for calculating the number of interconnects of a particular hierarchical level and the placement Rent's exponent for calculating the occupational probability of the interconnects in that hierarchical level. However, the usage of Rent's exponent is still only an approximation. Therefore, in [20] Dambre et al., improved upon the model developed by Stroobandt in [16] by eliminating the use of Rent exponent and thereby improving the accuracy in interconnect length distribution estimation.

The last decade has also seen a number of research efforts targeted at predicting the individual wire lengths. For example, the growth-limited multifold clustering methodology in [21] predicts the length of each net separately by performing a clustering based greedy local placement starting from each cell in the circuit. In [22], the post-routing net length is estimated from a polynomial function of several parameters for each net. The parameters are based on the local properties such as the number of pins in the net, number of nets of each unique pin counts in the neighborhood, and global parameters such as the number of cells and pins in the design and the place and route tool. Recently in [23] a mutual contraction (MC) metric has been developed to identify the shorter connections in the given netlist.

While the above explained methodologies were aimed at a priori interconnect length prediction, the methodology in [24] enables an online interconnect length estimation. In other words this prediction is done dynamically during the process of top-down partitioning based hierarchical placement procedure. Consequently, this method results in more accurate estimations for online estimation purposes. Each of the above methodologies explained in this section is aimed at predicting the length of connections or interconnects or nets in the circuit at different stages of the design cycle with varying inputs resulting in varying accuracies.

Proposed Research

In spite of the wide variety of length prediction models proposed and developed most of them suffer from a severe lack of accuracy. However, the ideal solution to most of the problems arising out of lack of information about the wires would be an accurate prediction of the exact individual post-routing wire lengths of all the wires in the design, during the early stages of the design flow. This would provide a good perspective on all the relevant issues at stake, such as the timing, delay, repeater planning, power dissipation, and heat dissipation attributes of a given design choice.

While predicting the individual routed wire length is ideal, it is a very difficult proposition since it will necessitate bringing into account routing level details into the prediction model. This could complicate the prediction process. Further, it will make it harder to identify possible sources of error in the prediction model. Besides, placement is the primary phase that establishes the basic lower limits on the wire length. Therefore, this research is limited to pre-routing, post-placement wire length prediction.

Most wire length prediction methods can be categorized into two types based on their objective. Microscopic models provide individual wire length prediction, while macroscopic models provide wire length distribution prediction. Another way to classify the wire length prediction models is a priori prediction and online prediction, where the former predicts the wire lengths before the physical design process, whereas the latter provides a continuously evolving wire length prediction during the hierarchical physical design procedure. The research in this work is limited to a priori prediction using both microscopic and macroscopic prediction models due to the larger scope of applicability. However, due to the large number of possible placement solutions, the lengths of the wires will vary from one solution to another. This will limit the accuracy of the prediction. Further, there are also limitations to prediction based on the modeling methodology. Consequently, these limitations to prediction are studied first, following which new models are developed.

Subsequently, this thesis is organized into six chapters. Following this introductory chapter, the first two chapters deal with macroscopic prediction, and the next two chapters deal with microscopic prediction. In each of these pairs, the first chapter deals with the study on limitations to the corresponding prediction (macroscopic or microscopic), while the second deals with the corresponding new model development. Finally, chapter six summarizes the conclusion and future work of this thesis.

The research in this thesis is carried out under the assumption that the individual cell (gate) size variations do not affect the quality of prediction by a huge margin. For the

sake of simplicity in model development and investigation of limitations, it is assumed that the cells (gates) in a given design are of unit size with unit gate pitch width and height. Therefore, IBM placement benchmarks [25] modified to have cells of unit size are used in the investigations. Further, the layout area of the benchmarks is also modified by limiting them to square shapes with approximately 5% white space. The input/output terminals are also removed from the netlists to limit the prediction to wires inside the chip. The attributes of the resulting benchmarks are shown in Table 1.1.

TABLE 1.1
BENCHMARKS USED IN THE RESEARCH AND THEIR ATTRIBUTES

Benchmark	Gate Count	Net Count	Interconnect Count
IBM01	12036	11507	28672
IBM02	19062	18429	52823
IBM03	21924	21621	50298
IBM04	26346	26163	58533
IBM05	28146	28446	79985
IBM06	32019	33354	85392
IBM07	44848	44394	108270
IBM08	50691	47944	130565
IBM09	51461	50393	123024
IBM10	66948	64227	176722
IBM11	68119	67016	154986
IBM12	69026	67739	190052
IBM13	81018	83806	196967
IBM14	145492	143202	339640
IBM15	157861	161196	423765
IBM16	181633	181188	499030
IBM17	182359	180684	540059
IBM18	210051	200565	539258

Finally, it should be noted that five different placement tools are used in this research to place these circuits. They are Dragon 3.01 [26], Capo8.8 [27], FengShui5.0 [28], mPL5.0 [29] and an In-house Simulated Annealing based Placement Tool. Partitioning is performed using hMetis [30] and the Rent exponent is extracted using the Rent exponent calculator Rentc [31].

CHAPTER 2

LIMITATIONS OF MACROSCOPIC PREDICTION

Macroscopic prediction of wire lengths have both inherent limitations and limitations arising out of the model used to make the prediction. A careful study of these limitations is necessary to understand the bounds to macroscopic prediction and to reduce the prediction error. Therefore, investigations were carried out to understand these limitations, and the results from the investigations are presented in this chapter.

The ideal solution to all wire related problems would be an accurate prediction of the length of every wire. But only a fraction of the wires have reasonably similar and therefore predictable length from one placement solution to another (see Chapter 4). Therefore, the next best solution to the wire problem is to provide an accurate macroscopic perspective by predicting the length distribution accurately. However even this could be difficult, since the length of the individual wires and subsequently their length distribution will vary from one placement solution to another. It will vary both from run-to-run within a placement tool and from tool-to-tool. For example, in [32] it was reported that nearly 30% variation in solution quality was observed in commercial placement and routing tools due to non-functional changes in the tool input such as renaming variables or permuting lines of gate level netlist. Also in [33] a similar study was performed that resulted in a difference of up to 7% between the best and the worst placement result.

Therefore, the first part of this chapter presents the results from a study of the inherent limitations to macroscopic prediction due to variability in the distributions from one placement tool to another and from one placement run to another. The second part of the chapter deals with the basics of the current wire length distribution models and the limitations to accuracy arising out of the model themselves.

Variability of Distributions

Interconnect length distribution models dominate the current macroscopic model research due to the fact that modeling net lengths in macroscopic models is very difficult. Accurate prediction of interconnect length entails an accurate understanding of the variability associated with it. Therefore, in this section the variability in the average interconnect length is studied first, and then the variability in interconnect length distribution is studied. For each of the above cases, both intra-tool and inter-tool variability are studied. The Intra-tool variability, which is the variability in results within a placement tool, is studied using five different placement results extracted from Dragon3.01. Although it is possible for the intra-tool variability to change form one tool to another, it is assumed that the results from Dragon3.01 are a sufficient approximation to provide a basic understanding of the intra-tool variability. Inter-tool variability, which is the variability in results from one placement tool to another, is also studied using five different placement results, but with one result each from five different placement tools viz., Dragon3.01, Capo8.8, FengShui5.0, In-house Simulated Annealing and finally mPL5.0 global placement combined with FengShui5.0 based detailed placement. The detailed placement of the mPL5.0 global placement is done using FengShui5.0 because the detailed placer of mPL5.0 crashed when used on the modified benchmarks used for our experiments.

Coefficient of Variation

The metric used in this study to analyze the variation in a given variable is called the coefficient of variation (CoV). It is defined as the ratio of the standard deviation of a given variable to the average of that given variable. CoV is used to study the variation because the metric isolates the variation from the magnitude of the average value of the data by providing the standard deviation per unit value of the average. This is especially useful in applications that compare variables with a large range of average values.

Average Interconnect Length

The average interconnect length is the most basic point measure of interconnect length, and its value, which is derived from interconnect length distribution is used in a variety of applications such as area estimation and power consumption [34]. Table 2.1 gives the average and coefficient-of-variation of the average interconnect length for both inter-tool and intra-tool variability for the various benchmarks. It was observed that for the intra-tool variability, the CoV has a range of 0.025 with an average of 0.0183. The CoV of the inter-tool variability has a wider range of 0.0672 and a greater average of 0.0815 that is nearly 4.5 times the intra-tool CoV. It can also be seen from these results that the average interconnect length for a given netlist has a worst case CoV of 0.0316 for intra-tool variability and 0.1186 for inter-tool variability.

TABLE 2.1

AVERAGE INTERCONNECT LENGTH INTRA-TOOL AND INTER-TOOL VARIABILITY

Benchmark	Intra-tool Variability		Inter-tool Variability	
	Average	Coefficient of Variation	Average	Coefficient of Variation
IBM01	7.72	0.0214	7.59	0.0619
IBM02	13.96	0.0188	13.75	0.0797
IBM03	12.98	0.0157	12.85	0.0750
IBM04	11.91	0.0092	11.92	0.0604
IBM05	22.38	0.0207	21.86	0.0514
IBM06	13.81	0.0316	13.63	0.1034
IBM07	12.69	0.0116	12.54	0.0899
IBM08	15.69	0.0241	16.13	0.0785
IBM09	10.78	0.0299	10.83	0.0719
IBM10	12.20	0.0163	12.35	0.0755
IBM11	11.14	0.0088	11.41	0.0764
IBM12	14.62	0.0219	14.92	0.0894
IBM13	12.21	0.0066	12.25	0.0766
IBM14	15.49	0.0151	15.65	0.0766
IBM15	15.07	0.0167	15.47	0.0878
IBM16	16.17	0.0300	16.43	0.0983
IBM17	18.98	0.0169	19.24	0.0957
IBM18	17.73	0.0142	18.50	0.1186
Average		0.0183		0.0815

Interconnect Length Distribution

In this section, the variation in the distribution of interconnect lengths is analyzed by calculating the CoV of the distribution for each length. Figure 2.1 shows a sample relation between CoV of the length distribution and length of interconnects of IBM03 benchmark for inter-tool variation. It can be observed from this figure that the coefficient of variation generally increases with length of the interconnect. Similar CoV relations were observed for the other benchmarks in both intra-tool and inter-tool variation study. Based on these observations it is possible to hypothesize that *the distribution of shorter interconnects with higher frequency are more predictable (repeatable) due to its lower CoV than the distribution of longer interconnect*s. Also it was observed that the CoV of the individual length distributions is much higher for most lengths than the CoV of the average interconnect length; for this reason it can be hypothesized that the prediction of the distribution of interconnects will be more difficult than the average length.
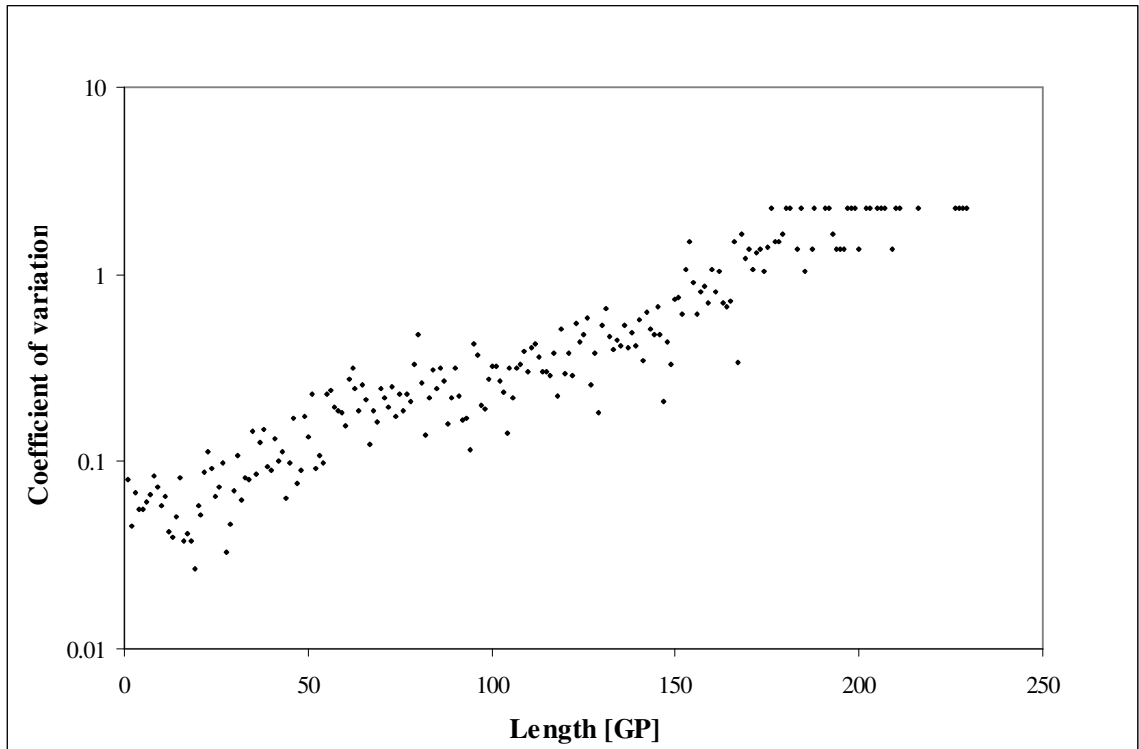
Figure 2.1 Inter-tool length distribution variability for IBM03

Further analysis is therefore performed to develop a better understanding of distribution variability. This is done by dividing the entire length range in the distribution into two regions, those lengths with less variation in the distribution (predictable region) and those lengths with high variation in the distribution (unpredictable region). A fixed CoV value is used as a threshold (repeatable threshold) to separate the two regions. The predictable part of the distribution is defined as that part of the distribution starting from length one and extending until the length for which the five-point moving average of the CoV of the length distribution exceeds the repeatable threshold. The five-point moving average was used to smoothen the relationship between the CoV of length distribution and the length. The value of repeatable threshold is varied from 0.1 to 0.5. This corresponds to a standard deviation that is less than 10% of its average distribution to less than 50% of its average distribution.
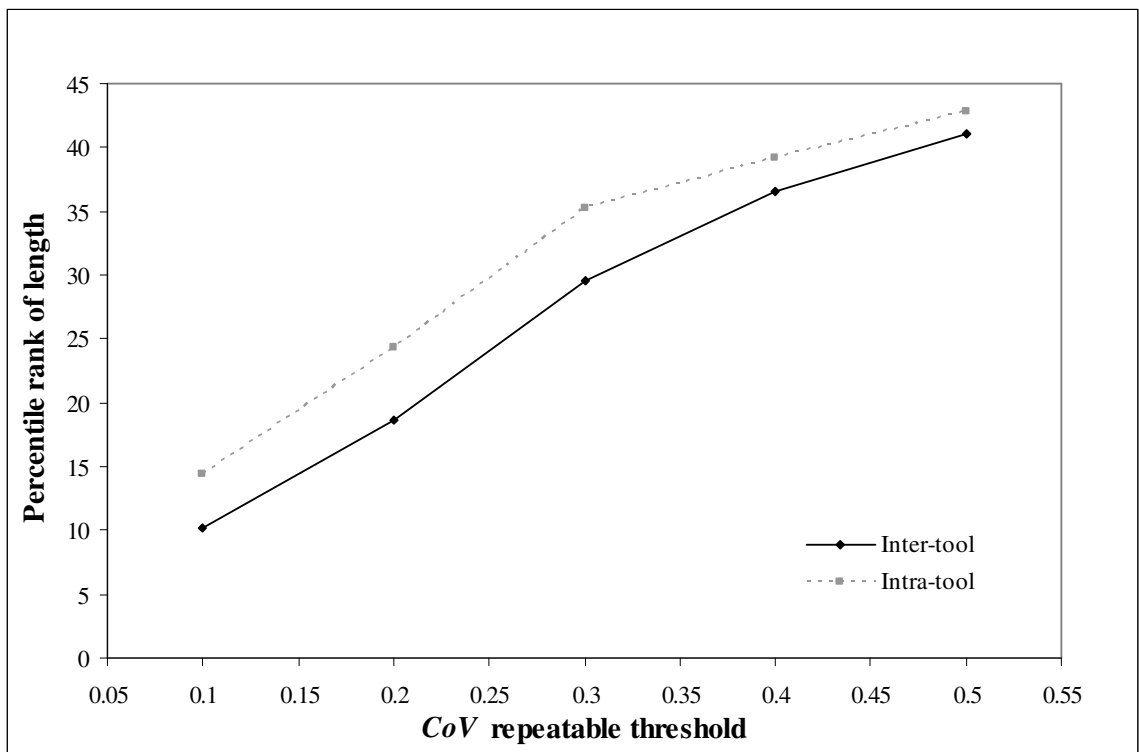
Figure 2.2 Percentile rank of length below which the distribution is repeatable

Figure 2.2 shows the relation between the repeatable threshold and the average of the percentile rank of length (from five different benchmarks IBM01 to IBM05) below which the distribution is repeatable for both intra-tool and inter-tool variation. Percentile rank of length is the percent value of the length with respect to the maximum possible length of an interconnect in the given design. On an average the shortest 14% of the lengths in a given design has a distribution value whose standard deviation is less than 10% of the average distribution, for the case of intra-tool variation (see Figure 2.2 for a CoV of 0.1). However for the case of inter-tool variation (see Figure 2.2) on an average only the shortest 10% of the interconnect lengths meet the same repeatable threshold criteria of 0.1.
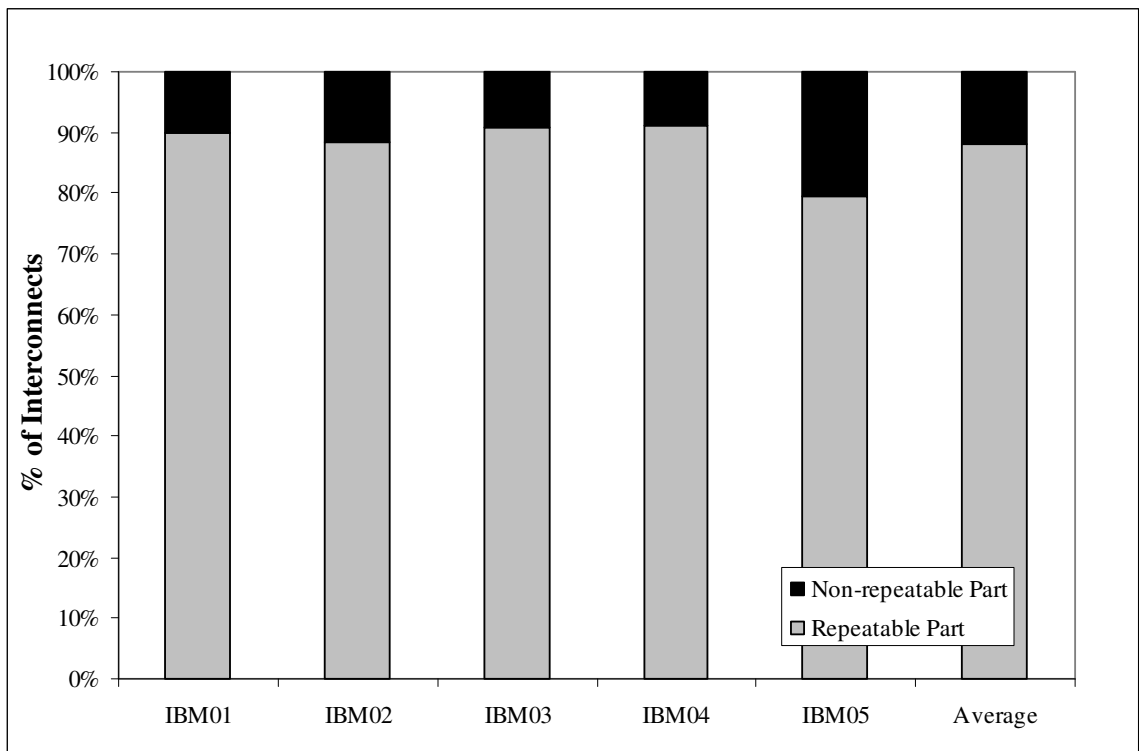
Figure 2.3 Percent interconnects in the repeatable part of distribution

An interesting attribute of this repeatable part of the distribution in the inter-tool variation is that this shortest 10% of interconnect lengths constitute 88% of the total

number of interconnects on an average. This forms the predictable region (repeatable part) of the distribution, and the remaining 12% of interconnects constitute the part of the distribution that varies widely (non-repeatable part) from placement tool-to-tool. This result depicting the relation between the percent interconnects in the repeatable and non-repeatable part of the distribution is shown in Figure 2.3 for five benchmark circuits.

Yet another interesting aspect of this repeatable part is that the 88% of interconnects in the repeatable part constitutes only 52% of the total interconnect length, while the remaining 12% of interconnects in the non-repeatable part constitute nearly 48% of the total interconnect length. This result in terms of the percent total interconnect length in the repeatable and non-repeatable part of a distribution is shown in Figure 2.4 for five benchmark circuits.
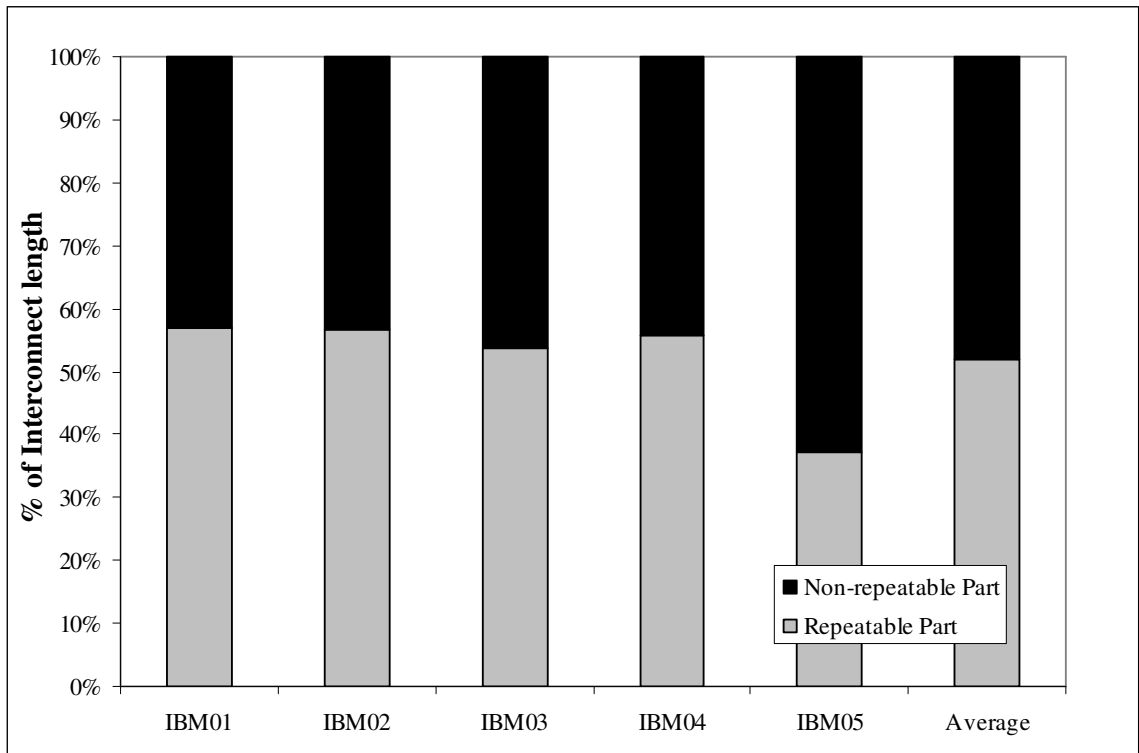
Figure 2.4 Percent interconnect length in the repeatable part of distribution

A typical application for wire length distribution such as repeater insertion planning would need the total wire length of the longer global interconnects and therefore

the distribution of the wires in the non-repeatable part of the distribution with longer wires. Table 2.2 gives the observed relationship between the best-case and worst-case scenario (minimum and maximum) of the percentage of total interconnect length in the non-repeatable part of the distribution among the set of five different distributions studied under both intra-tool and inter-tool variation. From this table it can be seen that the maximum of total length of interconnects in the non-repeatable part of the distribution could be 13% more than the minimum in the case of intra-tool variation, and 39% more than the minimum in the case of inter-tool tool variation. Therefore if a length distribution is predicted with discrete values, even if it is accurate for one run, it could still underestimate the total length of the interconnects in the non-repeatable part of the distribution by as much as 13% for another run of the same tool or as much as 39% of for another run in a different tool. Consequently *any repeater insertion planning done based on a discrete predicted interconnect length distribution without keeping these variations in consideration will suffer*.

TABLE 2.2
PERCENT TOTAL INTERCONNECT LENGTH IN NON-REPEATABLE PART OF DISTRIBUTION

Netlist	Inter-tool variation		Intra-tool variation	
	Minimum (% Length)	Maximum Minimum	Minimum (% Length)	Maximum Minimum
IBM01	27.37	1.18	39.58	1.32
IBM02	26.67	1.09	37.76	1.52
IBM03	41.55	1.07	42.76	1.43
IBM04	26.52	1.09	39.31	1.43
IBM05	37.05	1.20	59.27	1.25
Average	31.83	1.13	43.73	1.39

In such scenarios, a model to predict the variation would greatly aid the interconnect length distribution application development effort. Figure 2.5 shows the relationship between the average of the distribution values of interconnect length from five placement results and their corresponding CoV's observed for the five benchmark

circuits in the inter-tool variation study. A clear relationship between the CoV and distribution of interconnect lengths can be observed from this figure.
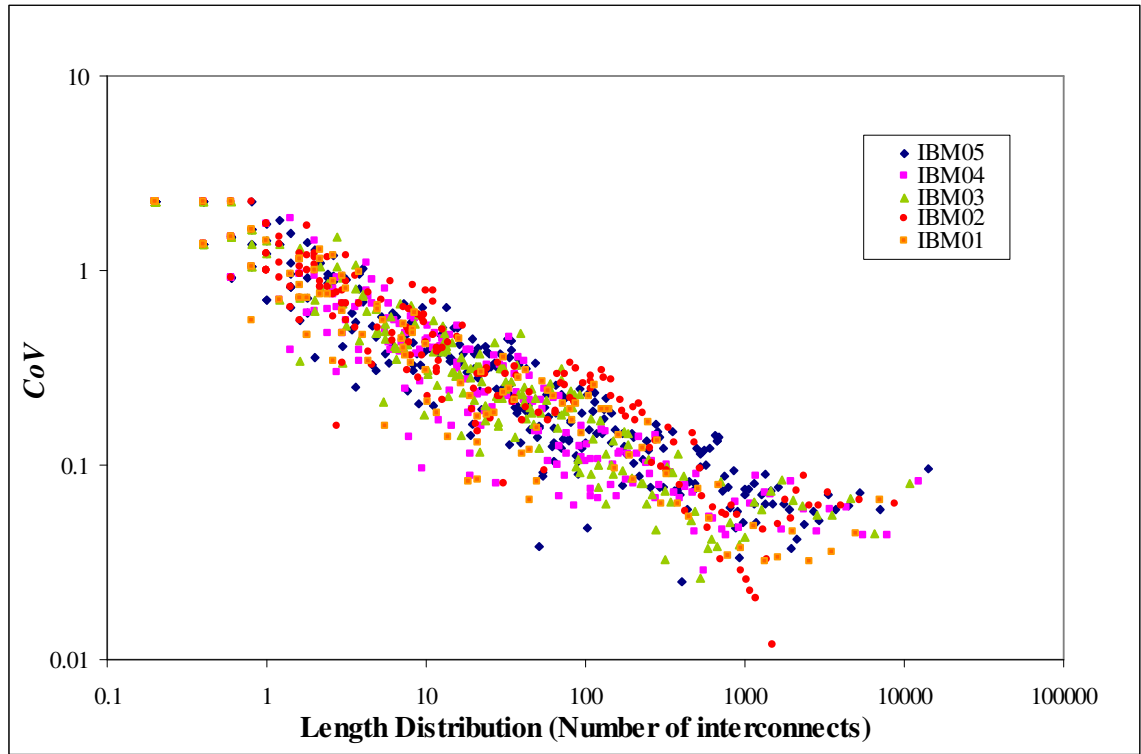
Figure 2.5 Relation between length distribution and its coefficient of variation

Consequently, a piecewise model is developed by dividing this data into two regions. Linear regression is then used to fit the logarithmic values of length distribution and its coefficient of variation. The resulting piecewise model is shown in Figure 2.6. For distribution values less than 1000, the model shows a strong negative correlation with an R^2 value of 0.9028, while for distribution values greater than 1000, which are observed for very short lengths, the correlation is very weak with the CoV values being almost constant. Although the model is empirical, it still depicts some basic relations between the variation of the length distribution and the distribution for a given length. Therefore, given a reasonably accurate discrete length distribution model approximate estimations on the variation of these distributions is possible.
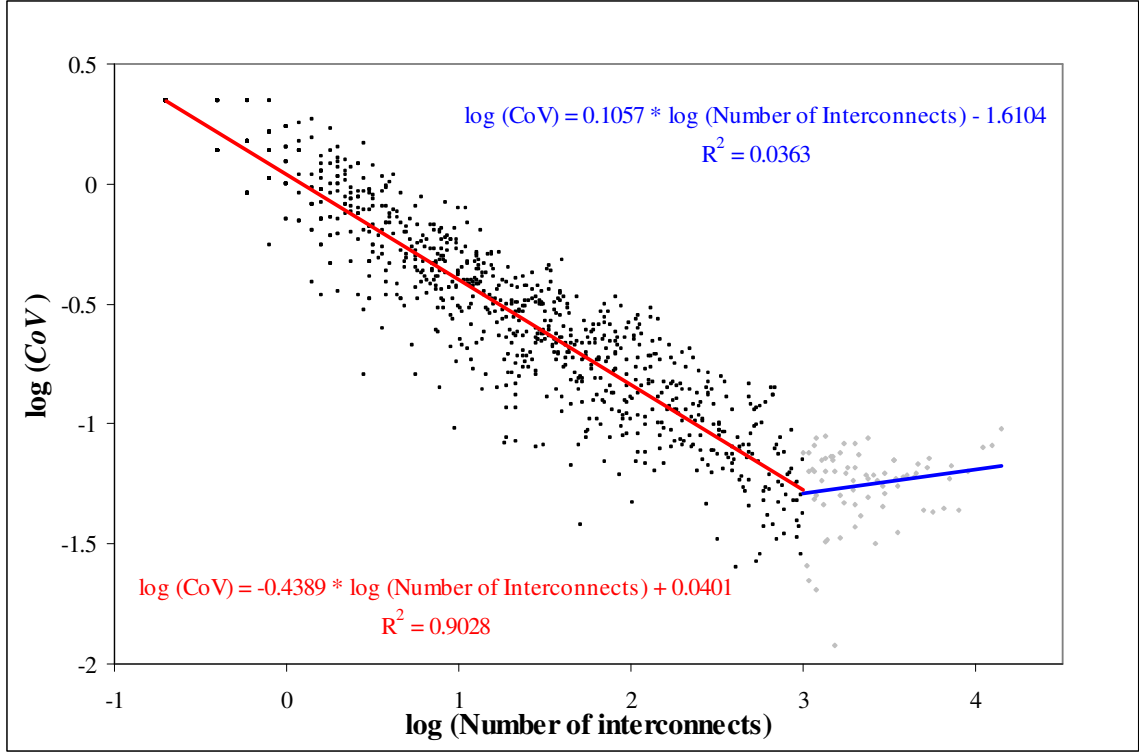
Figure 2.6 Piecewise model of the coefficient of variation of length distribution

Macroscopic Models Overview

Macroscopic models are mainly limited to interconnect length prediction due to the difficulty and time requirements involved in modeling the length of wires with more than two terminals as a net model in the prediction methodology. They generally model the length distribution of the wires. The earliest macroscopic model was developed by Donath [9]. Current state of the art macroscopic models were developed by Davis, De and Meindl (DDM model) [14], Stroobandt et al., (Stroobandt's model) [15] and an improved version of the later developed by Dambre (Dambre's model) [20]. All of the above three models are derived using a fundamentally common framework. They estimate the number of interconnects of a given length by using a combination of two functions. The first part called *interconnect site density function* calculates the number of interconnect sites of a given length available for the interconnects to occupy. The second

part called the *interconnect occupational probability function* gives the probability that an interconnect position of a given length is occupied.

The major difference between these models lies in the nature of the derivation of the solution. Stroobandt's model and Dambre's model are derived using Donath model as the basis. Since the physical design process can be viewed as a hierarchical process corresponding to the partitioning of the circuit followed by the partitioning of its corresponding layout in which the partitioned block will be placed, the models adopt a hierarchical approach, assuming the system will be placed using such a methodology. The model works by estimating the number of interconnects newly cut at each partitioning level. This is followed by an estimation of the number of interconnect sites of different lengths available for these newly cut interconnects in the correspondingly newly partitioned layout blocks at each hierarchical partitioning level. The interconnect positions are counted such that they only include those positions that have one terminal each in each of the layout partitions of the correspondingly newly partitioned layout. An interconnect occupational probability function is then used to estimate the probability that one of the newly cut wires occupy one of the available positions in each hierarchical level. Thus an interconnect length distribution is estimated for each hierarchical partitioning level and the total interconnect length distribution is obtained as the sum of the interconnect length distributions of the individual hierarchical levels. DDM model on the other hand does not make any assumptions about the placement methodology used and calculates the interconnect site density and occupational probability for the entire design in one step, in a flat manner without any hierarchical modeling. In this section of the chapter, a closer look is taken at the DDM model and Dambre's model to understand their limitations.

All of these models ignore the variation in cell sizes and assume all the cells to be of a uniform size. The DDM model and Stroobandt model assume these cells to be of unit size with a square shape. Dambre's model on the other hand includes the possibility to

model rectangular shaped cells [20]. The models assume that the cells are placed in a layout, which is in the form a grid with the sites for the cells having the same size and shape as the cells.

DDM model

The DDM model of interconnect length distribution for a given circuit is derived by using a *principle of conservation of terminals* between three blocks constructed over the layout. The three blocks A, B and C used in the development of the model, shown in Figure 2.7, have N_A, N_B and N_C cells in them respectively. The blocks are constructed such that a single closed path can encircle one, two or all three of the blocks and together form a semi-manhattan circle with the single celled block A at the centre, the cells of block C at the periphery, and the cells in between forming block B. Consequently the cells in block C are at a fixed distance from the cell in block A. This block based system is used to derive an expression for estimating the number of terminals between blocks A and C. This function is then used to derive the expression for interconnect occupational probability function $Ip[l]$, where l is the distance between the block A and block C. Finally the interconnect length distribution function is obtained by the combination of interconnect site density function $M[l]$ and interconnect occupational probability function $Ip[l]$.

According to the conservation of terminals, the terminals coming out of a block should either be connected to one of the other block's terminals or to one of the terminals that lies outside the blocks. Applying this principle to the three block system in Figure 2.7 yields Eq2.1, where T_A, T_B, T_C and T_{ABC} represent the total number of terminals coming out of block A, B, C and ABC and $T_{A\text{-}to\text{-}B}$, $T_{B\text{-}to\text{-}C}$ and $T_{A\text{-}to\text{-}C}$ represent the total number of terminals between blocks A and B, B and C and A and C respectively.

$$T_A + T_B + T_C = T_{A-to-C} + T_{B-to-C} + T_{A-to-B} + T_{ABC} \qquad \text{(Eq2.1)}$$
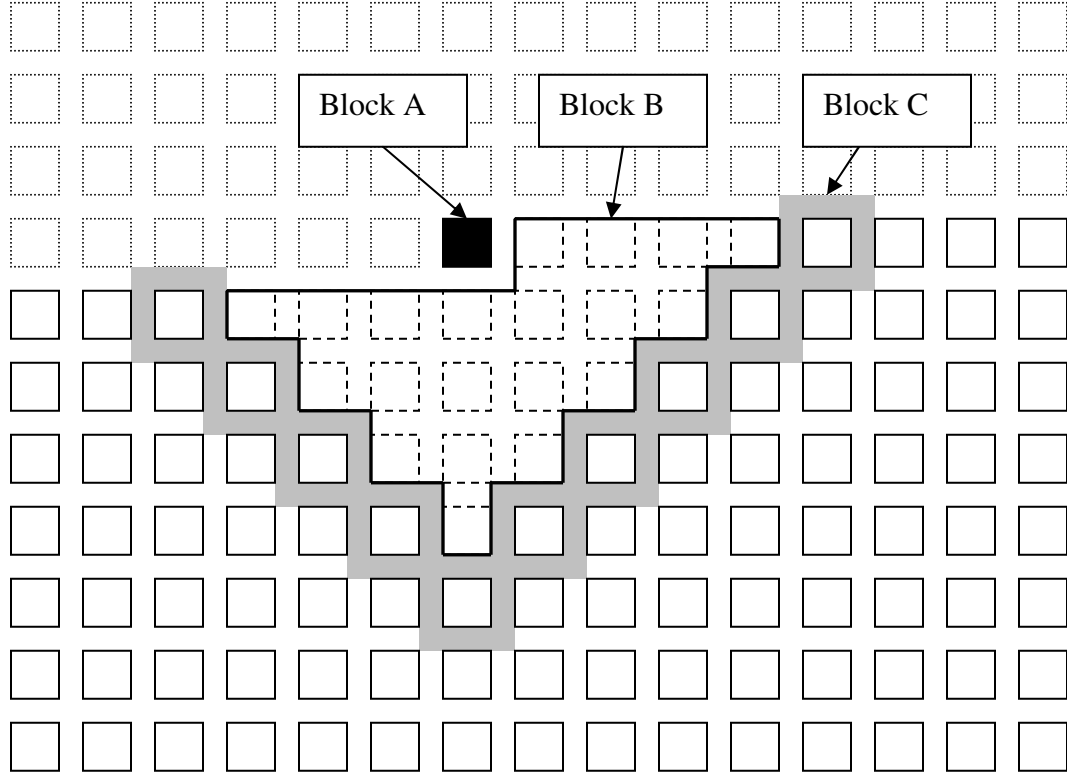
Figure 2.7 Blocks used in interconnect length distribution model derivation

Upon simplification Eq2.1 yields Eq2.2, which gives the total number of terminals between blocks A and C, where the cells in block C are separated from the cell in block A by a constant distance, say l.

$$T_{A-to-C} = T_{AB} - T_B + T_{BC} - T_{ABC} \qquad (Eq2.2)$$

The terminal count values used in Eq2.2 are based on the net models terminals. Assuming a single source terminal for each net, each sink terminal of the net will correspond to a single interconnect. Hence a factor α (Eq1.8), which gives the fraction of net terminals that are sink terminals is used to calculate the number of interconnects between block A and C (I_{A-to-C}), that are of length l by multiplying the number of terminals between block A and C by α.

$$I_{A-to-C} = \alpha \left(T_{AB} - T_B + T_{BC} - T_{ABC} \right) \qquad (Eq2.3)$$

The probability that there is an interconnect of length l ($Ip[l]$) in between a pair of cell positions separated by a distance l is given by the probability that there is an interconnect between one of the cells in block C and the cell in block A.

$$I_P[l] = \frac{\alpha(T_{AB} - T_B + T_{BC} - T_{ABC})}{N_C}$$

(Eq2.4)

The total number of interconnects of a given length l in the system ($I[l]$) is then obtained by the product of the probability that there is an interconnect in between a pair of cell positions separated by length l given by $Ip[l]$ and the total number of cell position pairs that are separated by a distance l given by $M[l]$ (Eq2.5). The expression for number of pairs of cell positions separated by a given distance l is derived by counting the number of cell positions in Block C for all possible positions of block A in the layout, and is given by (Eq2.6)

$$I[l] = M[l] \cdot Ip[l]$$

(Eq2.5)

for $1 \le l < \sqrt{C}$:

$$M[l] = \left(\frac{l^3}{3} - 2l^2\sqrt{C} + \frac{1}{3}l(6C - 1) \right)$$

for $\sqrt{C} \le l < \left(2\sqrt{C} - 2\right)$:

$$M[l] = \left(-\frac{l^3}{3} + 2l^2\sqrt{C} - \frac{1}{3}l(12C - 1) + \frac{2}{3}\sqrt{C}\left(2\sqrt{C} - 1\right)\left(2\sqrt{C} + 1\right) \right)$$

(Eq2.6)

To calculate $I[l]$ values, it is necessary to have a methodology to estimate the total number of terminals from blocks B, AB, BC and ABC required by Eq2.4. It is for this purpose that the Rent's rule is used to estimate the total number of terminals coming out of the blocks B, AB, BC and ABC as shown in Eq2.7.

$$T_B = k(N_B)^p$$
$$T_{AB} = k(N_A + N_B)^p$$
$$T_{BC} = k(N_B + N_C)^p$$
$$T_{ABC} = k(N_A + N_B + N_C)^p$$

(Eq2.7)

The original interpretation of the Rent's rule in [2] is derived using a terminal-to-gate relation (T-to-G) where the terminal count for a block with a given number of gates is counted based on the number of segments into which the net is cut among the various blocks, where each cut net segment has one terminal coming out of its block to connect to the other cut net segments. This type of interpretation of terminals will be referred to as *net-terminal interpretation* and is illustrated in Figure 2.8.

Net-terminal interpretation
Net source gate: p; Net sink gates: q, r, s
Block 1: p, r; Block 2: q, s; Block terminal count: 2 (shaded squares)
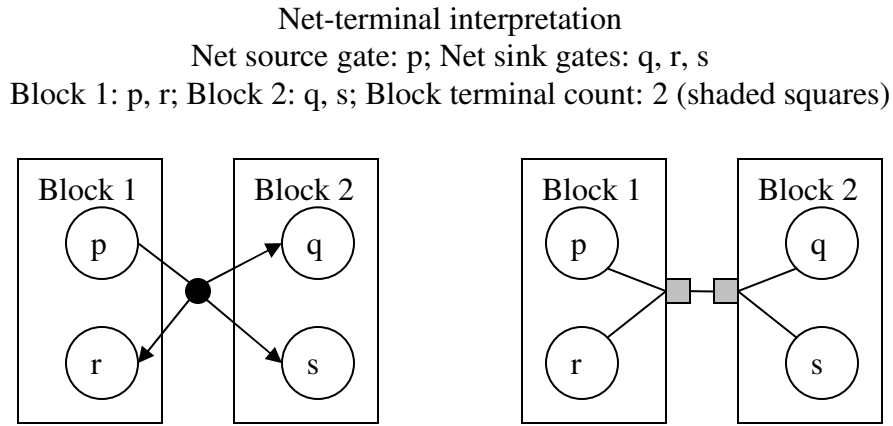
Figure 2.8 Net model (left) and net-terminal interpretation (right)

Given Eq2.7, values of the sizes of the blocks A, B and C are needed to estimate the required terminal count values. However the sizes of the cells will vary based on the position of block A within the layout. For example, if the block A in Figure 2.7 is placed close to the periphery of the layout and if the distance between block A and C that is considered is large enough, the semi manhattan circle forming the blocks may partially lie outside of the layout area. In order to make calculations simpler, the DDM model is derived assuming that the layout grid in which the cells are placed is infinitely large, and the block sizes can be approximated by assuming that the blocks B and C will always be entirely within the layout. Eq2.8 gives the expression for the size of the blocks based on this assumption. This approximation will be referred to as *infinite plane block size approximation*. Using Eq2.8 and Eq2.7, the expression for *Ip*[*l*] in Eq2.4 is simplified by binomial expansion to its final form shown in Eq2.9.

$$N_A = 1$$
$$N_B = l(l-1) \qquad \text{(Eq2.8)}$$
$$N_C = 2l$$

$$Ip[l] \cong \alpha k \frac{p}{2}(2-2p)l^{2p-4} \qquad \text{(Eq2.9)}$$

In [14], the distribution model is fitted to an actual distribution, and the corresponding Rent exponent p is extracted. This Rent exponent value is then used to predict the length distribution of other designs. The main application of this technique in [14] was to predict wire length distributions for future generic systems to understand the limits of Moore's law. However, for applications where a rapid estimation of a wire length distribution is needed for a specific netlist, a different methodology is used by researchers, which is shown below.

Step 1: Extract the net-terminal interpretation based T-to-BS relation by performing partitioning of the netlist.

Step 2: Use the T-to-G relation to get the Rent-exponent (p) or use a generic p ~ 0.6 (when netlist information is not available for partitioning).

Step 3: Use Rent-exponent to calculate interconnect occupational probability and thus the length distribution.

Step 4: Renormalize the distribution to match the total number of interconnects in the system.

Dambre Model

Dambre's interconnect length prediction model [20] is derived using Donath's model as the basis. First, it explicitly converts all the wires with more than two terminals into two terminal interconnect model based wires. Then, the model relaxes some of the constraints of the Donath's model. First, Dambre's model allows the cells to be of a rectangular shape where the Donath's model limited them to a square shape. Second, Dambre's model allows the layout to take a rectangular shape, where Donath's model

limited it to a square shape with a further restriction that the number of gates in the layout be a power of 4.

In order to accommodate these architectural relaxations, Dambre's model allows the block sizes to be non-uniform at different partitioning steps. Consequently, it does not strictly adhere to a hierarchical methodology, explained in the introduction of this section, to calculate the interconnect length distribution. Instead, it works by maintaining a list of partitioned layout module sizes with their frequency and partitioning the largest module from this list at each stage and calculating the interconnect length distribution of the cut interconnects for each such partitioning step. Initially this list contains the entire layout as a single module, and the partitioning is carried on until all the modules are of the size of a single cell.

At each partitioning stage of this process, the partitioning of the layout module is done according to a set of cut rules to yield rectangular blocks. After each partitioning step, the number of interconnects that is newly cut at that level is calculated using the principle of conservation of terminals, which yields Eq2.10.

$$N_{cut} = \frac{T_A + T_B - T_{AB}}{2} \qquad \text{(Eq2.10)}$$

In Eq2.10 A and B are the two modules obtained by partitioning a larger module AB. The model then uses a *scaled* T'-to-G' relation that accounts for the white space in actual layouts, instead of the actual partitioning based T-to-G relation, to estimate the terminal counts for blocks A, B and AB. This scaled T'-to-G' relation is obtained from a partitioning based T-to-G relation by assuming that the white space will be uniformly distributed across the layout, and therefore among the different partitions. The expression used to transform this partitioning T-to-G relation into scaled T'-to-G' relation is given by Eq2.11, where T(g) is the number of terminals of a block with g gates in the partitioning T-to-G relation.

$$T'\left(\left(\frac{G'}{G}\right)g\right) = T(g) \qquad \text{(Eq2.11)}$$

Then, given the sizes of the partitioned modules A and B, the model estimates the number of interconnect sites that exist in between the gate positions in the two modules using an interconnect site density function. Expressions to estimate this inter-module site density are developed using generating polynomials and is explained in [35] by Stroobandt and Marck.

And finally the model adapts an occupational probability function proposed by Verplaetse in [18] to calculate the probability that one of the cut interconnects occupy one of the interconnect sites available in between the partitioned layout modules A and B. If the module AB is cut horizontally, then the magnitude of its vertical dimension Y that is being cut is used to estimate the occupational probability as shown in Eq2.12.

$$Ip[l,Y] \cong \frac{1}{1 + \left(\dfrac{l}{(Y/2)}\right)^{4-2p}} \qquad \text{(Eq2.12)}$$

It should be noted that as the length increases, the second term in the denominator of Eq2.12 starts to dominate and consequently it approaches the occupational probability values of Davis model and Stroobandt's model. The later two models use an occupational probability function with the dependence shown in Eq2.13.

$$Ip[l] \propto l^{2p-4} \qquad \text{(Eq2.13)}$$

The interconnect length distribution of the cut interconnects is then calculated using the inter-module site density values calculated for the partitioned modules and the interconnect occupational probability values calculated from Eq2.12.

However, to estimate the length distribution, a T-to-G relation extracted from actual partitioning is still necessary. This relation is necessary both to estimate the Rent exponent p value, which is used to estimate occupational probability, and to estimate the number of interconnects cut at each level of partitioning; however, partitioning takes

time. Therefore, a parametric model called β-model is suggested to quickly estimate the partitioning T-to-G relation. By using the β-model it is not necessary to partition the circuit completely. Instead it is sufficient to partition only a few levels and use the model to generate the relation for the smaller partitions. In fact, Dambre, Stroobandt and Campenhout argue that 4 levels of partitioning values are sufficient for generating a T-to-G relation that is accurate enough for estimating the wire length distribution in [20].

The average interconnect length results estimated from this model have been shown to exhibit good correlation to actual average interconnect lengths. However, this correlation was observed for a netlist that was placed after modifying its structure, by replacing the nets with more than two terminals by nets of two terminals, based on the interconnect model of a wire. But in reality, the netlist is placed and optimized without modifying the structure of the wires.

In any case, it can be observed that both of the above explained models (DDM model and Dambre's model) uses an estimate of Rent's exponent to calculate the occupational probability and therefore fundamentally similar to each other. But, the pseudo-hierarchical nature of the Dambre's model adds layers of complexity to the model. While this layered approach helps reduce the error with respect to Donath's model, it also leads to a difficulty in accurately pointing sources of error in the length distribution estimation process. On the other hand, DDM model is much simpler and therefore easier to analyze. Further the DDM model is more widely used in interconnect length distribution applications such as power estimation [4], thermal modeling [5], timing estimation [34], chip size estimation [34] and estimation of wiring demand and routability [36]. Therefore in the following subsection the results obtained using the DDM model is carefully evaluated.

Evaluation of DDM Model Results

The first step in the estimation of the DDM model interconnect length distribution is the extraction of T-to-G relation. Figure 2.9 shows the net-terminal interpretation based T-to-G relation for the IBM01 benchmark circuit and the corresponding Rent's rule with rent-coefficient (4.755) and rent-exponent (0.5371). The values of the Rent parameters are obtained by fitting a power law equation to a portion of the T-to-BS relation identified as Region I of the curve. This Region I is identified by the shaded region in Figure 2.9 and indicates the data points used to extract the Rent's parameters.
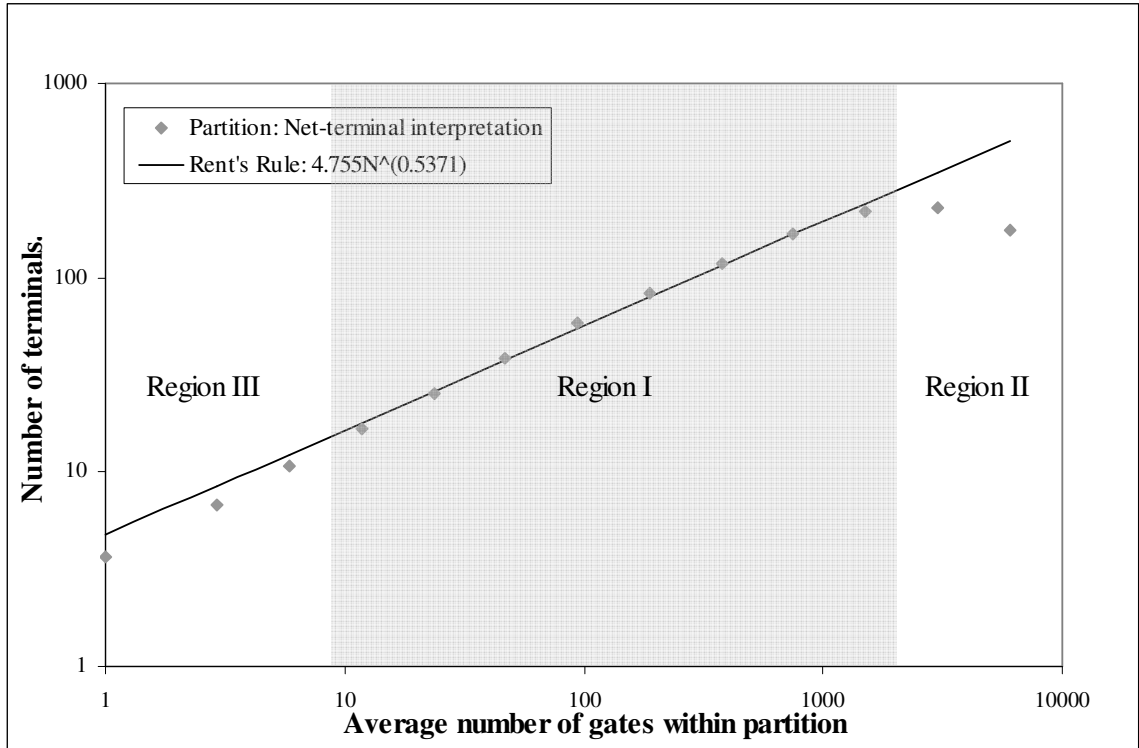
Figure 2.9 Terminal-to-Gate relation and the Rent's rule for IBM01

The Region I is characterized by an almost constant exponent in the power law relation required to model the terminal count for the corresponding block sizes. This Region I typically does not include the region of the curve for very large and very small block sizes due to the fact that a single power law relation do not fit well with these two extremities. The region of the curve corresponding to the larger block sizes is referred to

as the Region II of the curve and the region corresponding to the smaller block sizes of the T-to-G relation is referred to as Region III of the T-to-G relation.

Using the Rent's parameters so measured, the interconnect length distribution for the IBM01 circuit is calculated using the methodology outlined earlier in this chapter. The resulting distribution is shown in Figure 2.10 along with the actual interconnect length distribution extracted from placement results of three placement tools (In-house-simulated-annealing, Dragon3.01, Capo8.8).
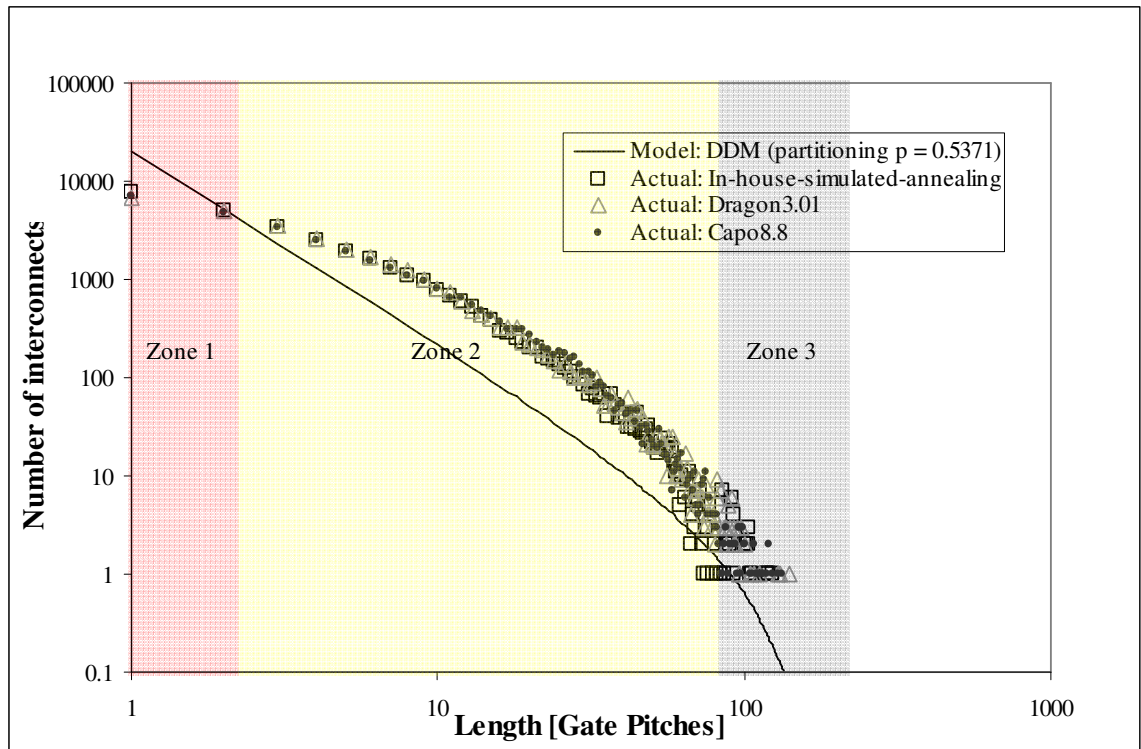
Figure 2.10 Actual and model interconnect length distribution for IBM01

It can be observed from this figure that the model distribution follows a certain trend when compared to the actual interconnect length distribution. Based on this inter-relationship the distribution can be separated into three non-overlapping continuous zones of lengths. Zone 1 contains the distribution of shortest interconnects, where the model distribution unmistakably overestimates the actual distribution. Zone 2 spans the short to medium length range of the distribution, where the model recognizably underestimates the actual distribution. Zone 3 comprises the distribution of longest interconnect lengths,

where the model neither overestimates nor underestimates in a clear manner. The three regions can be characterized by the following attributes, viz, zone starting length (StartL), zone ending length (EndL), percent interconnects in the zone for actual placement distribution (%IC), percent error in the cumulative distribution of the zone in model distribution with respect to the cumulative distribution of the zone in actual distribution (%ErrorCD). Similar trends were observed between the model and actual interconnect length distributions on other benchmark circuits as well, and the values of these zone attributes for the DDM model length distribution with respect to the interconnect length distribution of in-house-simulated-annealing placement results is shown in table 2.3 for five benchmark circuits.

TABLE 2.3
ZONE ATTRIBUTES OF DDM MODEL LENGTH DISTRIBUTION

Zone	Circuit	StartL	EndL	%IC*	%ErrorCD*
1	IBM01	1	2	39.42	+91.77
	IBM02	1	3	31.74	+113.48
	IBM03	1	3	41.99	+71.79
	IBM04	1	2	34.20	+85.88
	IBM05	1	4	33.65	+109.07
	Average	**1**	**2.8**	**36.21**	**+94.39**
2	IBM01	3	66	60.33	-59.86
	IBM02	4	74	67.04	-53.39
	IBM03	4	118	57.67	-52.01
	IBM04	3	128	65.56	-44.72
	IBM05	5	203	66.26	-55.33
	Average	3.8	117.8	**63.37**	**-53.06**
3	IBM01	67	218	0.25	-25.68
	IBM02	75	275	1.20	-19.22
	IBM03	119	295	0.32	-45.85
	IBM04	129	323	0.23	-22.20
	IBM05	204	335	0.07	-63.66
	Average	118.8	289.4	**0.42**	**-35.14**

* * %IC: % Interconnect count (cumulative distribution) in zone; %ErrorCD: % Error in cumulative distribution between model results and actual simulated annealing based placement results in a zone

From Table 2.3 it can be observed that slightly more than one third of the interconnects fall within zone 1 which is made up of the very short interconnects.

However, the model, on an average, overestimates the number of interconnects in this zone to twice its original value. Consequently the distribution of the remaining two thirds of the interconnects in zone 2 and zone 3 is underestimated by a factor of 53.06% and 35.14% respectively.

It is interesting to note that ideally zone 1 of the distribution must be derived using the Region III of the T-to-G relation. Therefore any accurate modeling of this region of the T-to-G relation and using it directly in modeling the distribution could reduce the nearly 100% over estimation in the zone 1 of the distribution. Consequently, this must also reduce the error in the other zones of the distribution.

Since the distribution is calculated as a product of the occupational probability and the site density, and since the site density function has been shown to be accurate in [14], the above observed errors must arise from the occupational probably estimates. Therefore in the following section the various limitations, to modeling the distribution accurately, that arise from within the occupational probability function are examined.

Limitations of Macroscopic Models

For a macroscopic model such as the DDM model to be accurate, the following conditions must be met: (1) since it needs to calculate the *interconnect* length distribution, it needs to use the T-to-G relation in which the terminals for a given block size are counted based on the cut interconnect segments between blocks and not the cut net segments; (2) since it needs to calculate the distribution at the end of *placement*, the T-to-G relation used needs to be based on the placed layout and not based on an optimized partition; (3) since it calculates the distribution for each length based on conservation of terminals between the blocks B, AB, BC and ABC, it needs the exact terminal count for the four blocks B, AB, BC, ABC for each length and not the Rent's Rule based approximation; (4) since it calculates the terminal count based on the size of the block, it needs an accurate estimate of the average size of the blocks corresponding to

various lengths under the finite layout size constraint and not the approximations based on the infinitely large layout size approximation. Each of these limitations is examined in greater detail in the following subsections.

Terminal interpretation (Net vs. Interconnect)

The interconnect based terminal interpretation required to accurately model the interconnect length distribution is illustrated in Figure 2.11 and will be referred to as *interconnect-terminal interpretation*.

Interconnect-terminal interpretation:
Net source gate: p; Net sink gates: q, r, s; Interconnects: (p, q), (p, r), (p, s)
Block 1: p, r; Block 2: q, s; Block terminal count: 4

Figure 2.11 Interconnect model (left) and interconnect-terminal interpretation (right)

The problem with using net-terminal interpretation is that not only does the DDM model not need the net-terminal based terminal count, the conservation of the terminals also does not hold true for net-terminal interpretation. For the principle to hold true, each terminal must be connected to only one other terminal. It cannot be connected to more than one terminal as in net-terminal interpretation. An example illustrating the conservation of terminals principle with the two terminal interpretation methods is shown in Figure 2.12. Unlike DDM model, Dambre's model is not expected to suffer due to the terminal interpretation method because the model is applied to a netlist only after explicitly converting nets with more than two terminals into two terminal interconnects.

Net-terminal interpretation vs. Interconnect-terminal interpretation:
Net source gate: p; Net sink gates: q, r; Interconnects: (p, q), (p, r)

From figure above,
$T_p = 1$; $T_r = 1$; $T_{p\text{-to-r}} = 2$; $T_{pr} = 1$
Principle of conservation of terminals
=> $T_{p\text{-to-r}} = T_p + T_r - T_{pr}$
   L.H.S = 2;
   R.H.S = 1+1-1 = 1;
   L.H.S $\neq$ R.H.S
=> **For net-terminal interpretation**
   $\mathbf{T_{p\text{-to-r}} \neq T_p + T_r - T_{pr}}$

From figure above,
$T_p = 2$; $T_r = 1$; $T_{p\text{-to-r}} = 2$; $T_{pr} = 1$
Principle of conservation of terminals
=> $T_{p\text{-to-r}} = T_p + T_r - T_{pr}$
   L.H.S = 2;
   R.H.S = 2+1-1 = 2;
   L.H.S = R.H.S
=> For interconnect-terminal
interpretation $T_{p\text{-to-r}} = T_p + T_r - T_{pr}$
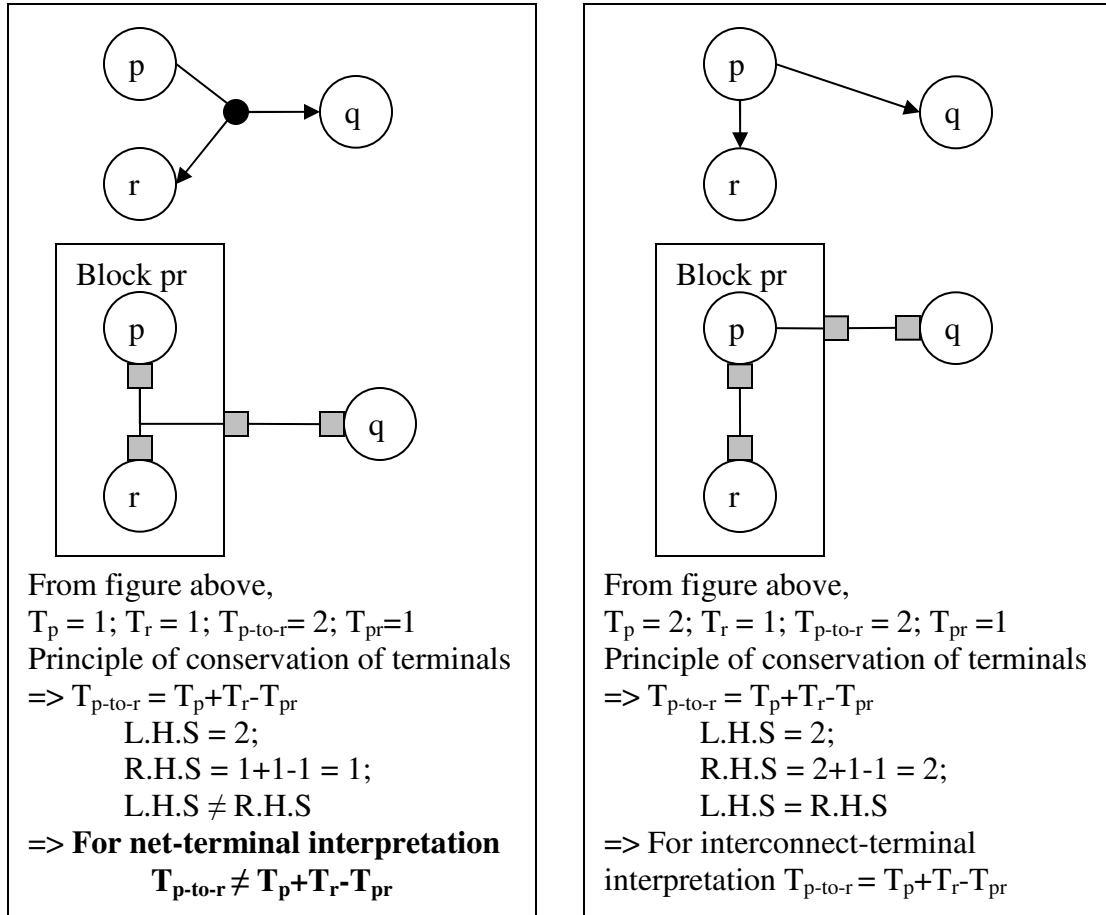
Figure 2.12 Validity of conservation of terminals

Physical design information (Partition vs. Placement)

Figure 2.13 shows the Terminal-to-Gate (T-to-G) relation based on both interconnect-terminal interpretation and net-terminal interpretation for partitioning results of IBM01 benchmark circuit. It also depicts a Terminal-to-Block Size (T-to-BS) relation based on interconnect-terminal interpretation of placement result for the same circuit. The relation is depicted with respect to block size because the DDM model requires the terminal count for a give block size in the placement, which is based on the total number of *gate positions* within the block and not merely the number of gates within the block.
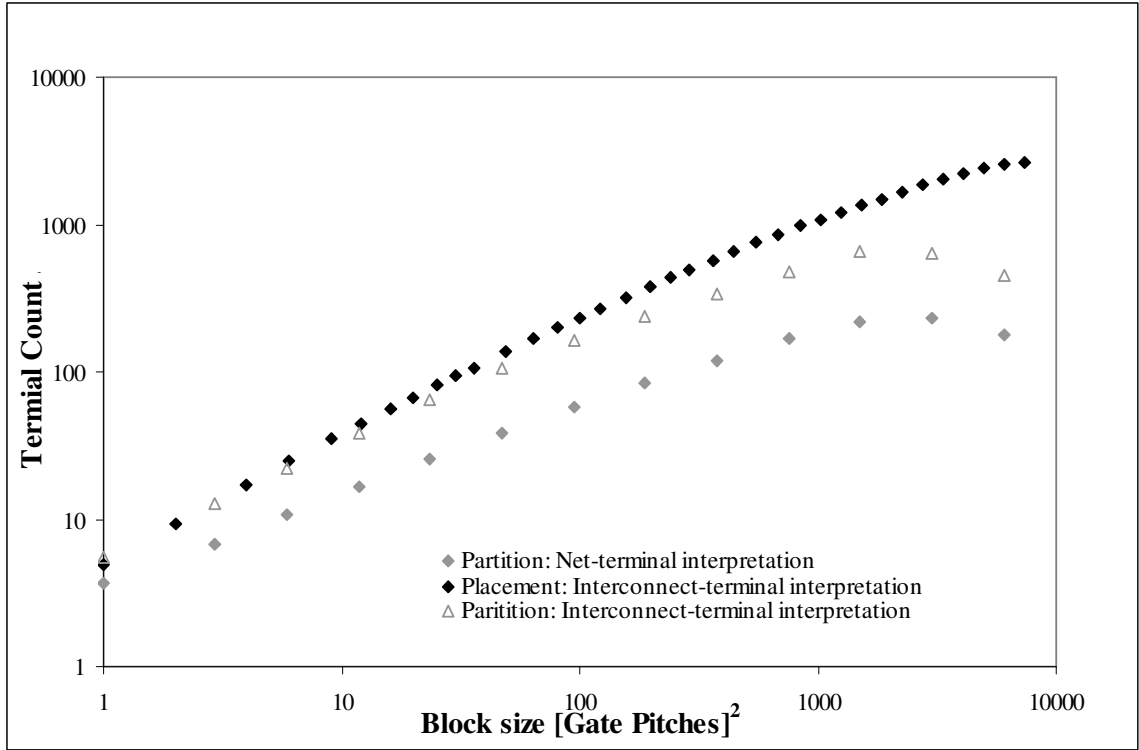
Figure 2.13 Terminal-to-Gate and Terminal-to-Block Size relation for IBM01

In this figure, the difference between plots *Partition: Net-terminal interpretation* and *Partition: Interconnect-terminal interpretation* reveals the impact of the terminal interpretation methods singularly on the T-to-G relationship. Partition based relation is extracted from partitioning the netlist into blocks that are non-overlapping in nature (see Figure 2.14 (a)). In partition, each cell in the design is assigned to one of the blocks at a given hierarchical level and the number of terminal coming out of these blocks is minimized.

(a) From partitioning   (b) From placement   (c) From placement
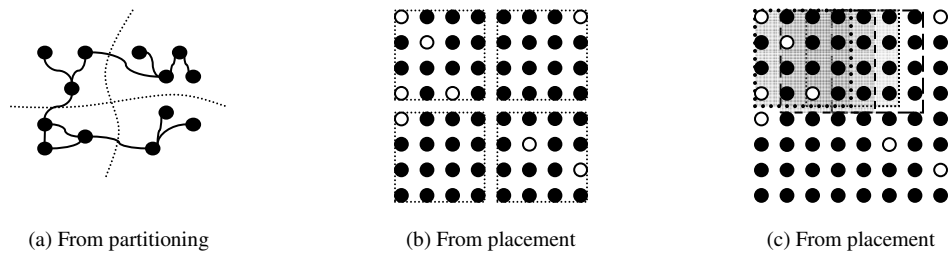
Figure 2.14 Terminal count extraction strategies

The DDM model requires the average number of terminals per block over all of its possible positions in a placement. It should not be extracted by dividing the placement into non overlapping hierarchical blocks as shown in Figure 2.14(b). Therefore, the placement based T-to-BS relation shown in Figure 2.13 is extracted by overlaying a rectangular block, of a given size and aspect ratio closest to one, on all possible positions within the layout of placed cells, and counting the number of terminals coming out of the overlaid block (see Figure 2.14 (c)). The placement of the overlaid block over all possible position results in several suboptimal block formation (in terms of number of terminals coming out of the block) when compared to partitioning. Further, the placement optimization causes an increase in the number of terminals coming out of a block because the placement cost function has to minimize the distance between the terminals of different blocks, when the blocks are placed on a two dimensional layout. While the above two factors tends to increase the average terminal count for a given block size, the presence of white space in the actual layout causes the average terminal count to drop, since it reduces the number of terminals coming out of a block. Consequently, there is a difference between *Partition: Interconnect-terminal interpretation* and *Placement: Interconnect-terminal interpretation* in Figure 2.13, and the use of partition based T-to-G relation instead of placement based T-to-BS relation introduces significant error into the model. This subject of different T-to-G relations is also treated in [18] and [19].

Terminal count (Actual Terminal-to-Block Size relation vs. Rent's Rule)

Figure 2.15 shows the error in terminal count for a given block size when the Rent's rule is used instead of the actual placement based interconnect-terminal-to-block-size relation for IBM01 benchmark circuit. It can be seen that the absolute error consistently increases with block size. It is clear therefore that the model in fact uses a largely erroneous terminal count by using the Rent's rule in place of the required placement based interconnect-terminal-to-block-size relation.
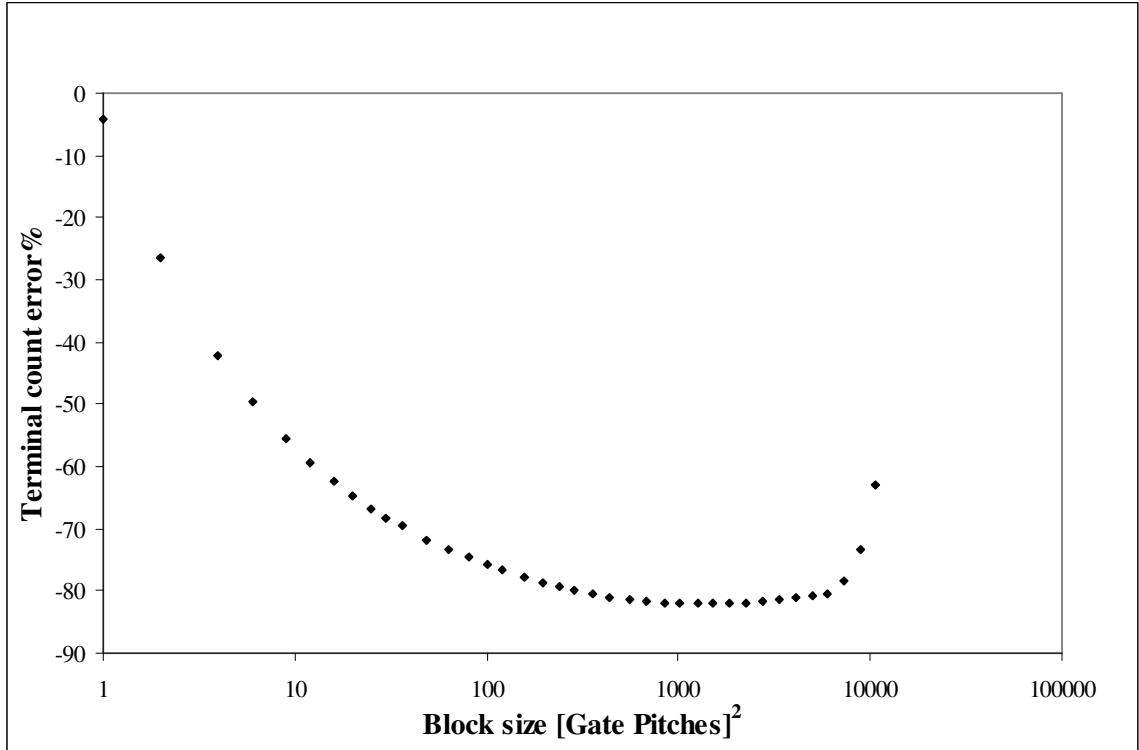
Figure 2.15 Error in terminal count as a result of using Rent's rule for IBM01

Block size (Infinite layout size vs. Finite layout size)

The DDM model needs the block size to estimate the number of terminal count of the block. But, the size of the blocks in the model varies with the distance between blocks A and C, where the distance is the same as the interconnect length whose distribution is estimated. For each such length, the basic shapes used in the model are similar to that of block B (see Fig 2.7). The DDM model assumes that for any position of the block A, all of the four blocks B, AB, BC, ABC used by the model will be completely within the layout area. However, in reality this is not true, since as the length increases and when the block A is closer to the periphery, only part of the block will fall within the layout and hence the actual average area of a block for each length will be lesser than the area of the blocks used by the DDM model. Figure 2.16 shows the error in the size of the block B with respect to the distance between blocks A and C. The error is the percent difference between the block size from infinite layout size approximation and the actual finite layout

49

size based block size. It is clear from this plot that there is a significant error in the block size used by the model, and that it increases continuously with distance between blocks A and C.
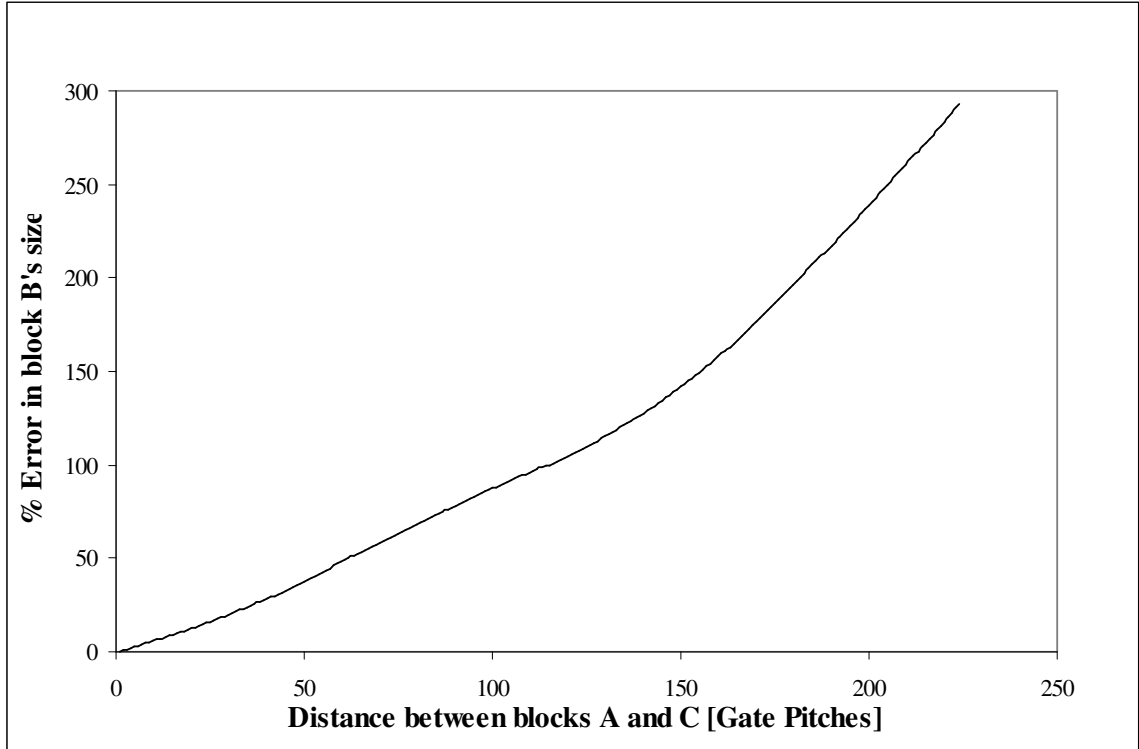
Figure 2.16 Error in average block size of block B for IBM01

Summary

Wire length varies from one placement solution to another. As a result there is variability in the wire length distribution. This could consequently be a limiting factor to achievable prediction accuracy. The intra-tool variability is generally less than the inter-tool variability. Further, the variability for shorter wires is lesser than that of longer wires, and that the variability increases with a decrease in distribution. However, most often the longer wires are the sources of bottlenecks in a design. Therefore, when developing applications using predictions of longer wires, such as repeater planning it is necessary to keep this variability into account. To aid in such scenarios, an empirical model that provides an idea of the variability of the distribution is also developed.

Accuracy of macroscopic wire length prediction could also be limited due to the modeling methodology used. Most of the current state-of-the-art wire length distribution prediction models calculate the distribution as a product of two functions viz., site density function and interconnect occupational probability function. The interconnect occupational probability values are most often calculated from a Terminal-to-Gate relation provided by the Rent's rule. The model developed by Davis, De and Meindl (DDM) is one of the very popular wire length distribution models. Investigations were carried out to study the limitations to prediction accuracy due to the methodologies used in this model. If Rent's exponent is extracted using a partitioning method, it was shown that the DDM model overestimates the distribution of very short wires, normally constituting 36% of the wires, by 94% on an average. Consequently, the distributions of the remaining wires were found to be underestimated. In addition to partitioning Rent's exponent, further investigations also revealed the use of Rent's rule approximation and incorrect block sizes used in deriving the terminal counts as potential sources of errors.

CHAPTER 3

NEW MACROSCOPIC MODEL

On-chip interconnect's increasing influence on the circuit performance and design time have been well documented. The primary interconnect attribute that is the source of this influence is its length. Therefore, to understand the macroscopic impact of interconnects on the overall system design, length distributions of interconnects have been studied and models to predict them have been developed [9] [12] [14] [15] [16] [20]. But the length distribution results from these models are characterized by a lack of correlation with results from individual circuits [37]. Consequently wire length distribution models are not as widely adopted for optimization of current designs as for technology extrapolation applications such as GTX [38] [39], BACPAC [40] AND GENESYS [41]. An accurate estimate of the interconnect length distribution can be used in a number of ways during the process of design planning and design optimization, such as for (i) estimating the wiring demand and routablity, [36] (ii) estimating the power consumption [4], (iii) estimating the number of repeaters needed in the given design [4], (iv) estimating the thermal requirements [5], (v) estimating clock cycle [42] etc...

Therefore, based on the analysis of limitations to macroscopic models presented in the previous chapter, a new interconnect length distribution prediction model was developed as part of this research. The new model is developed based on the Davis, De and Meindl (DDM) model [14] because it is widely used in a number of applications [4] [5] [43] [44] [45] [46] [47] [48]. The initial part of this chapter is devoted to the theory behind this new model, while the results produced by the model are presented towards the end of the chapter. Because all of the identified limitations of the DDM model are related to the Terminal-to-Gate relation used, the fundamental improvements are made possible from fixing these limitations and are explained in the following section.

Terminal-to-Block Size Relation

The objective of DDM model is to predict accurately the interconnect length distribution at the end of placement. Therefore it needs a T-to-G relation that reflects the placement results. Further, the relation should give the number of interconnect terminals for a *block of given size* in the placement. Since the block may include both gates and empty gate positions, a Terminal-to-Block Size (T-to-BS) relation is required instead of a Terminal-to-Gate (T-to-G) relation. In order to model the T-to-BS relation, the actual T-to-BS relation is first extracted and then a model is developed for this relation. A sample T-to-BS relation extracted from a placement using the actual basic block shapes of B[*l*] and AB[*l*] of the DDM model is given in Figure 3.1. The figure also shows the Rent's rule based models for the IBM01 benchmark circuit.
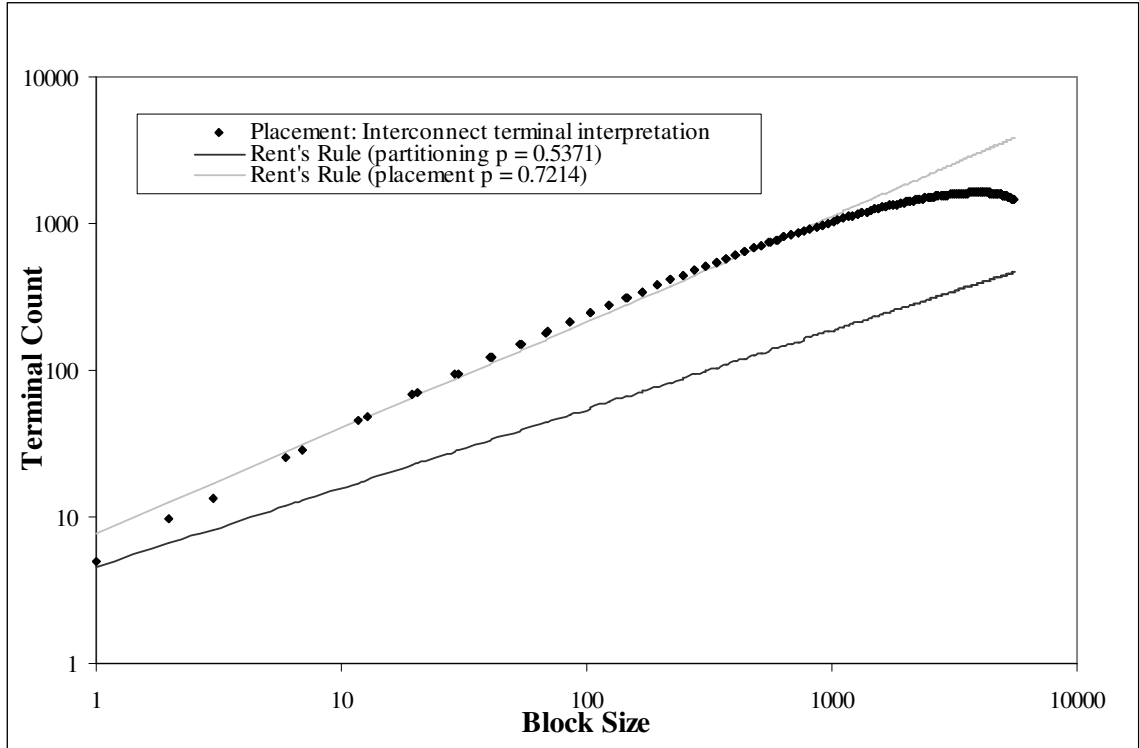
Figure 3.1 Actual and Model T-to-BS relation of IBM01

For a block comprised of a single cell position, based on Rent's rule, the number of terminals from the block is equal to the Rent's coefficient k.

$$T[1] = k(1)^p = k \qquad\qquad \text{(Eq3.1)}$$

Given a netlist with *I* interconnects that has to be placed in a square layout array with $N_{cellposition}$ number of cell positions, the average number of terminals per cell position is given by Eq.3.2

$$T[1] = \frac{2I}{N_{cellposition}} \qquad\qquad \text{(Eq3.2)}$$

Combining Eq.3.1 and Eq3.2 yields an expression for the parameter k as shown in Eq3.3.

$$k = \frac{2I}{N_{cellposition}} \qquad\qquad \text{(Eq3.3)}$$

The expression for Rent's rule can then be manipulated to provide Eq3.4. This is an expression for calculating the actual value of the exponent p for a given block size S (the number of cell positions within the block), given the terminal count for the block size and the value of k, where the value of k can be calculated from Eq3.3.

$$p[S] = \frac{\log\left(\dfrac{T[S]}{k}\right)}{\log(S)} \qquad\qquad \text{(Eq3.4)}$$

Block shape impact on exponent

Now given a T-to-BS relation, the actual p values can be calculated as a function of the block size. For example, two such relations are extracted using two different T-to-BS relations of IBM01 benchmark circuit and is shown in the Figure 3.2. The two different T-to-BS relations shown in Figure 2.13 and Figure 3.1 (as *Placement: Interconnect-terminal interpretation*) were extracted from the simulated annealing based placement result of the IBM01 benchmark using interconnect terminal interpretation. The result in Figure 2.13 is based on rectangular blocks (see Figure 2.14 (c)) and Figure 3.1 is based on actual block shapes B and AB (see Figure 2.7). The difference between the two

T-to-BS relationships is difficult to observe directly in a Terminal count vs. Block size graph. However, the difference is clearer in the Exponent vs. Block size graph.
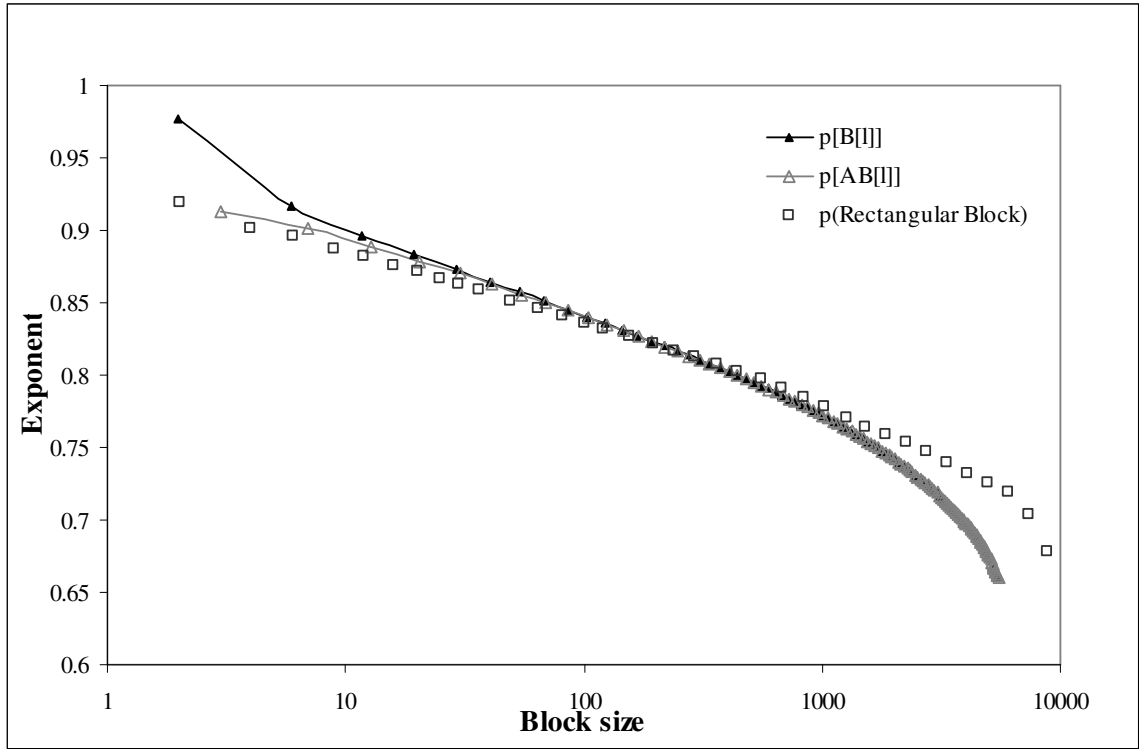
Figure 3.2 Actual exponent values for various block sizes of IBM01

It can be observed from Figure 3.2 that these p[S] values extracted from the placement are much higher than the Rent exponent value of 0.548 extracted by curve fitting the partitioning T-to-G relation of IBM01 benchmark. It can also be observed that they p[S] values vary with the block size. Furthermore, it can also be seen that for very small block sizes the p values of the actual block B follow a different trend when compared to that of block AB and the rectangular blocks. As seen in Figure3.3, this difference is due to the fact that, for very small block sizes, the cells are not tightly packed within the block B when compared to the block AB and the rectangular blocks of similar size. When the cells are loosely packed in the form of odd shapes(i.e., semi manhattan circles) to form a block as in the case of block B, the distances between the cells are greater than in a tightly packed block. Consequently, due to placement

optimization the rate at which the terminals are absorbed is lesser. As a result, the factor p, which is an indicator of the number of terminals coming out of the block, is *greater* than a block that is much tightly packed where the cell positions are closer. However, as the block sizes increase this difference reduces because most of the cells are surrounded by the cells within the block, and therefore the terminals are absorbed more or less the same. However for very large block sizes, blocks B and AB have lesser p values than the rectangular blocks. This is because the average terminal count is calculated over all possible positions of the basic block shape over the layout. In the case of rectangular blocks, all of its positions remain within the layout (see Figure 2.14 (c)). However, for positions with block A close to the periphery of the layout, as shown in Figure 3.3 only part of the block B fall inside the layout. Further as the length l used to characterize the blocks B and C increases, the number of positions at which only part of block B[l] and AB[l] remains within the layout increases. Consequently, in comparison to rectangular blocks, the average block size decreases and with it the average terminal count decreases as well. Similar relations between the three p values where observed on the other benchmarks as well. Given this relationship between the exponent p and block sizes, it is necessary to develop a model for the exponent.
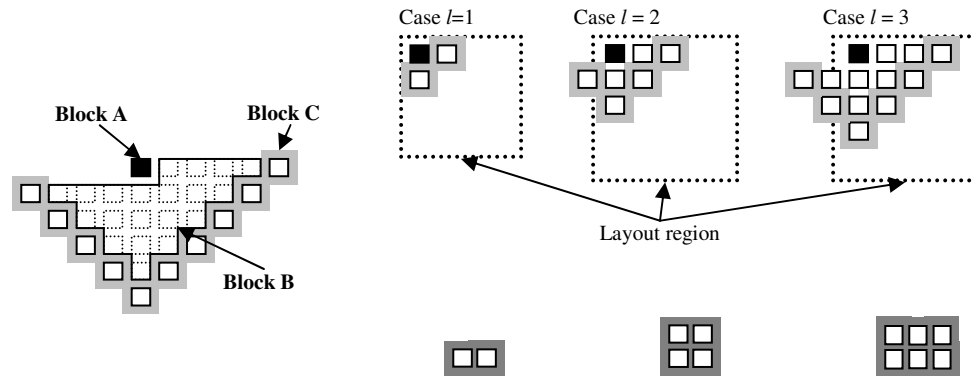
Figure 3.3 Actual block shapes used by model and rectangular blocks of small sizes

Exponent Model

According to the differential equation based interpretation of Rent's Rule, *p* is the fraction that reflects the level of placement optimization [49] such that when a new cell is added to a block with *S* cells and *T* terminals, the *rate of change of terminals* is given by *p*(*T*/*S*). From this, the parameter *p* can be interpreted as the fraction of terminals that are not absorbed inside the block when a cell is added to a block X of *S* cells, and so the fraction given by the parameter *p* can also be viewed as the *unabsorbed terminals factor*. Subsequently, the fraction (1-*p*) can be viewed to as the *absorbed terminal factor*, referring to the fraction of terminals that are absorbed within the block from the single cell added to the block.
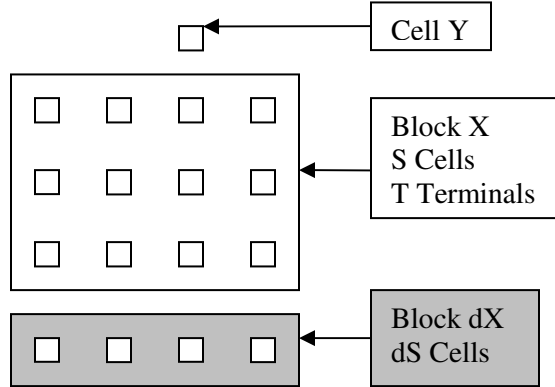
Figure 3.4. Blocks used to derive exponent p model

The relation between the parameter *p* and *S* can be understood by carefully examining the possible causes for the change in *p* (*dp*) with respect to an incremental change in the size of the block (*dS*). From Figure 3.4, let X be the block with *S* cells and Y be a cell adjacent to X such that the inclusion of Y to X will be characterized by a terminal increase of *p*(*T*/*S*). High *p* value indicates that a large fraction of terminals of cell Y are connected to the cells outside block X and vice versa.

Therefore, if the factor *p* is large, an incremental block dX of size *dS* adjacent to X will have an increased probability of absorbing these unabsorbed terminals because of

the presence of a larger number of unabsorbed terminals by the original block X. However, if the factor p is small reflecting a smaller fraction of unabsorbed terminals by block X, the probability of absorbing these terminals by the incremental block dX will be less. Therefore, the change in the absorbed terminal factor $(1-p)$ with respect to the incremental block dX of size dS will vary directly proportional to the unabsorbed terminal factor p, or

$$\frac{d(1-p)}{dS} \propto p \Rightarrow \frac{dp}{dS} \propto -p \qquad \text{(Eq3.5)}$$

As the size of block X increases, the cells of the incremental block dX will tend to be at a farther distance from the cell Y. But at the same time, due to placement optimization, most of the terminals of Y will be absorbed by the cells that are closer to Y and therefore fewer terminals will be left to be absorbed by cells that are far from cell Y. Accordingly, the rate at which the terminals of Y are absorbed by dX decreases as the distance increases. As a result the rate at which the terminals are not absorbed into dX also decreases. Therefore, the change in *unabsorbed terminals factor p* with respect to incremental change in X will have an inverse relation to the distance from the cell Y to the cells of block dX, which varies directly in proportion to the block size S, or

$$distance(Y \ to \ dX) \propto S \qquad \text{(Eq3.6)}$$

$$\frac{dp}{dS} \propto \frac{1}{distance(Y \ to \ dX)} \qquad \text{(Eq3.7)}$$

Combining equations Eq3.5, Eq3.6 and Eq3.7 gives

$$\frac{dp}{dS} \propto -\frac{p}{S} \qquad \text{(Eq3.8)}$$

A proportionality constant q is introduced in the above equation to yield Eq3.9.

$$\frac{dp}{dS} = -\frac{qp}{S} \qquad \text{(Eq3.9)}$$

The above equation can be solved for p and is given by Eq3.10.

$$p = jS^{-q} \qquad \text{(Eq3.10)}$$

58

This model is fitted to actual *p* values of rectangular blocks of IBM01 benchmark and the result is shown in Figure 3.5. Since the *p* values of the blocks ABC are closer to the p values of the rectangular blocks than that of block BC, as observed from Figure 3.2, Eq 3.10 is used to directly model p[ABC[*l*]]. The values for p[BC[*l*]] are then estimated from these values
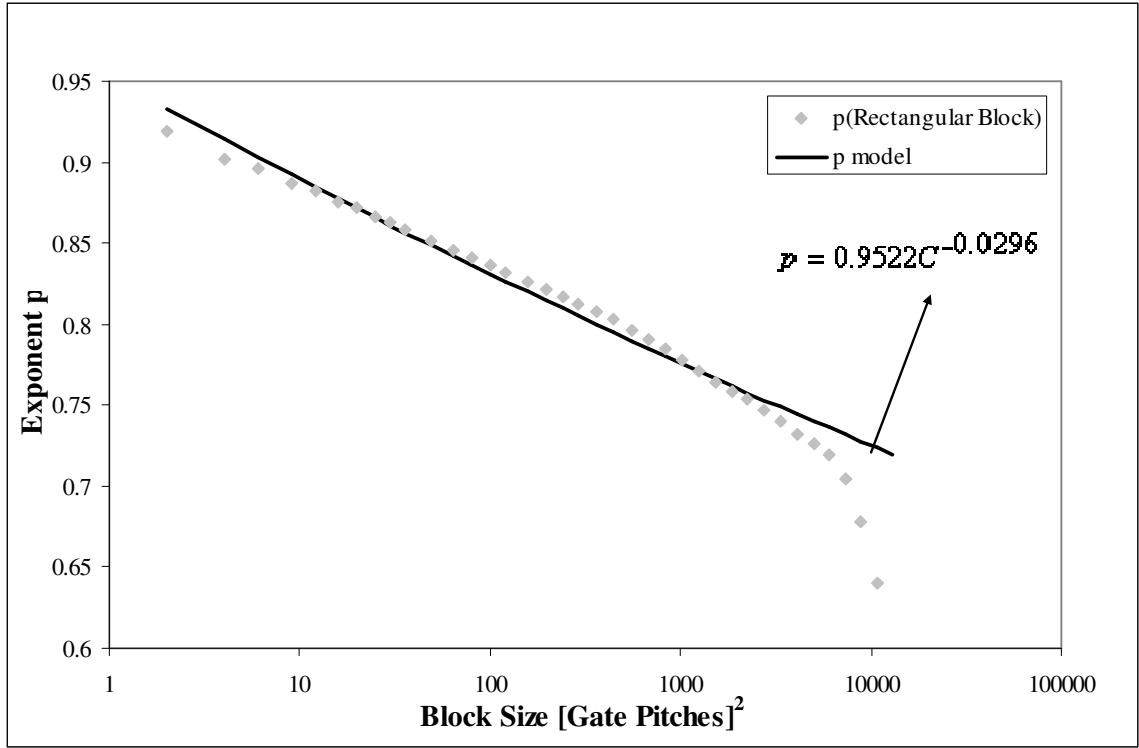
Figure 3.5 Model of exponent p fitted to actual p values of rectangular blocks

Exponent Model Parameters

Given the model for exponent *p*, it is necessary to develop a methodology to extract the values of the parameters of the model (*j*, *q*). This is done using a four step process outlined below and explained in detailed in the following subsections.

1. The value of the parameter *q* is estimated empirically.

2. A lookup table based model is developed to estimate a pair of exponent values *p*[B[*l*=2]] and *p*[AB[*l*=2]] referred to as the *initial exponent pair*.

3. The value of parameter j is obtained using Eq3.10 together with $p[AB[l=2]]$ and q to yield the p model for block ABC, $p[ABC[l]]$.

4. The values of p for block BC $p[BC[l]]$ is obtained using $p[ABC[l]]$ and $p[B[l=2]]$.

Parameter q extraction

The q values are extracted for the smallest nine benchmark circuits by fitting Eq3.10 to the relation between the factor p and the block size C extracted for rectangular block shapes. The rectangular block shapes are used because, for the most part, the values of p for the rectangular block shapes are very close to the values of p for the actual block shapes as shown in Figure 3.2. The corrections to take care of the deviation in p values for very small sizes of the actual blocks will be made using alternate strategies explained in the following subsections. The resulting q values have an average of 0.0253 and a standard deviation of 0.0036.   Further, they also exhibit a moderate correlation (correlation coefficient ~ 0.664) with the partitioning based Rent exponent calculated using publicly available Rent exponent calculator [31]. To avoid confusion between the new interpretation of p as a function of block size and the conventional Rent exponent interpretation with a single value for p, the later will be referred to as p_{rent}.

The plot depicting the relation between parameter q and p_{rent} with a linear curve fitted to this data is shown in Figure 3.6. With the lack of any other information, the average q value or the value derived by taking advantage of parameter q's correlation with the Rent's exponent p_{rent} (Eq3.11) could be used as an estimate of the actual q value.

$$q = \begin{cases} 0.0253 & \text{(average value model)} \\ & or \\ 0.0502 - 0.0411 p_{rent} & \text{(linear regression model)} \end{cases} \tag{Eq3.11}$$
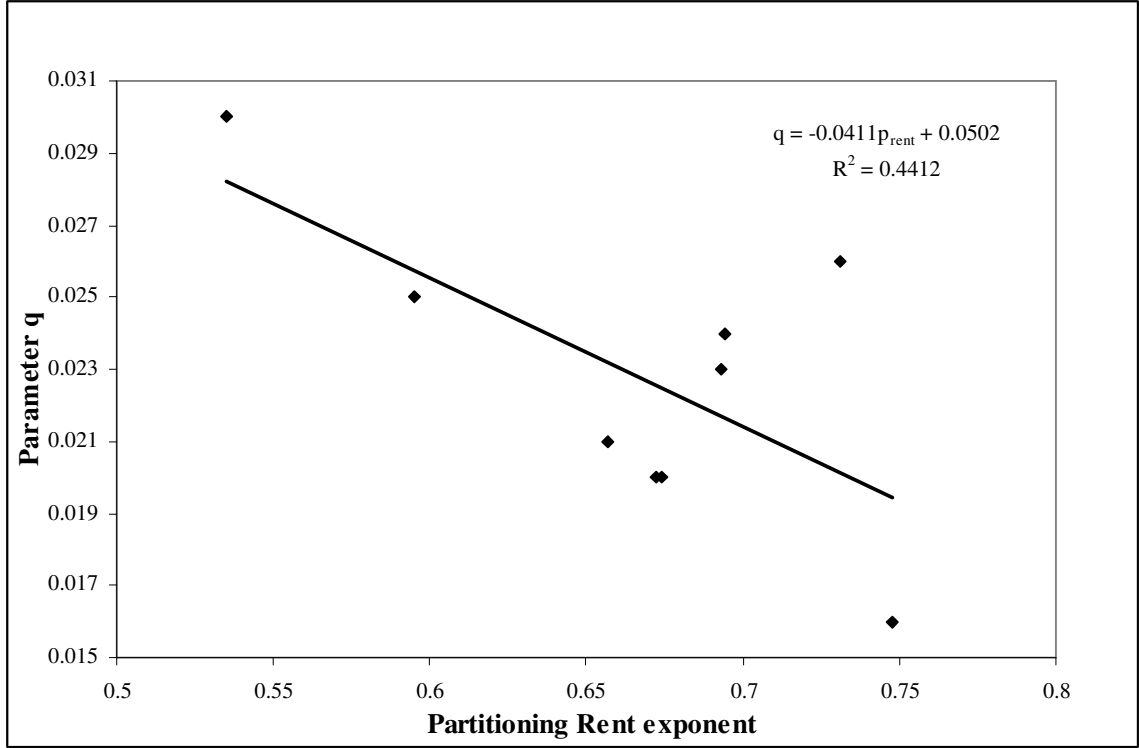
Figure 3.6 Parameter *q* vs. partitioning Rent exponent using IBM01-IBM09

Initial exponent pair model

It is the hypothesis of this research that the upper and lower limit of *p* values for the smallest blocks B[*l*=2] and AB[*l*=2] can be extracted from synthetic trees that share a single attribute with a real logic graph. This attribute is the number of interconnects per cell. It is believed that this is possible because the maximum number of very short interconnects in a graph is limited by the local nature of the graph and the limitations of the placement layout architecture. Furthermore, once the p-values for these small blocks are known, they can be used in conjunction with Eq3.11 to calculate j in Eq3.10.

For example, consider a simple tree graph formed by a reference gate with only its first level neighbor (Figure 3.7 (a)) and another that includes its second level neighbors as well (Figure 3.7 (b) and (c)). Assume that each gate, except the peripheral gates, is connected to the same number of gates and is equal to the average number of interconnects connected to a gate in a netlist (e.g. four).

61
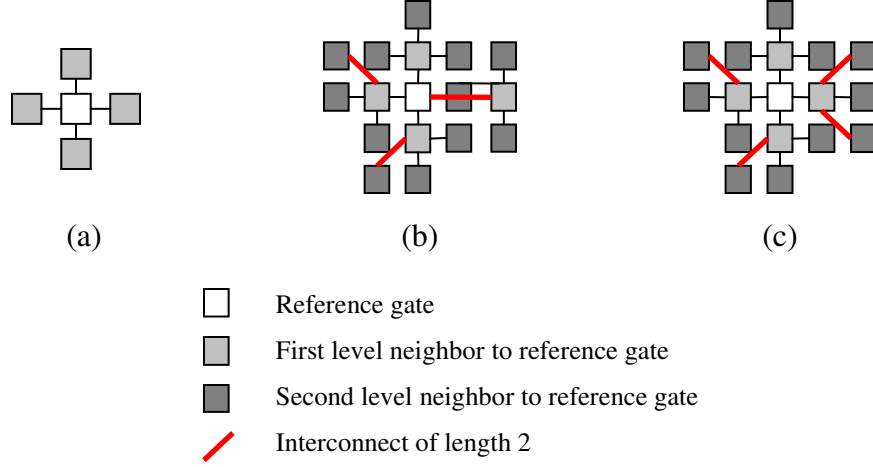
Figure 3.7 Influence of layout architecture and netlist structure on interconnects

Figure 3.7 shows three possible placement configurations for these graphs. Figure 3.7 (a) shows the optimal configuration for the reference gate and its first level neighbors with all interconnects being of length one. This arrangement can be obtained by greedily placing all the gates closest to each other. Figure 3.7 (b) shows a *greedy placement* with the second level of neighbors included; resulting in four interconnects of length two gate pitches each. Finally, Figure 3.7 (c) shows a better placement configuration for the same graph, with only three interconnects of length two gate pitches. Based on these, it can be said that the proportion of the shortest interconnects (l=1) is limited in the later cases (3.7 (b) and (c)) because (i) in 2-dimensional *placement layout architecture* there are only so many empty positions adjacent to the first level gates for the second level gate to occupy and (ii) the number of *neighbors at a given level of the graph* grows faster than the number of sites available, in this case at a rate of $4 \cdot 3^{d-1}$, where d is the number of levels of neighbors in the graph As a result, as the graph grows, a smaller proportion of the cells could be placed closer to each other and more cells have to be placed farther away resulting in longer interconnects.

From Figures 3.7 (b) and (c), it is clear that a measure of the occupation probability for length one or the corresponding p values $p[BC[l=1]]$ and $p[ABC[l=1]]$ extracted based on the greedy method would be suboptimal. Here, it should be noted that

$p[BC[l=1]]$ is the same as $p[B[l=2]]$, and $p[ABC[l=1]]$ is the same as $p[AB[l=2]]$ (see Figure 3.3). Therefore, such a mini-graph (local graph) placed using a *greedy strategy* would give an *approximate upper bound of p* for the blocks $BC[l=1]$ and $ABC[l=1]$ (see Figure 3.3 case $l=1$). However, due to the fact that the local graph is small and because it ignores many of the global netlist details, when placed using an *effective optimization strategy such as Dragon* the local graph would give the upper bound of the proportion of interconnects of length one, and therefore the *lower bound of p* for the blocks $BC[l=1]$ and $ABC[l=1]$.

Based on the above theory, tree based graphs were generated for different values of the average number of interconnects per cell I_{avg} (four, five, six, seven and eight), and placed using a greedy method and Dragon3.01. The number of levels of neighbors in the graph starting from a reference gate was set to either three or four so that the number of gates in the graph was at least hundred. This minimum limit of hundred cells was established empirically to ensure that the graph will be a reasonable representation of actual netlists. Each of the gates, excluding the gates at the periphery, where connected to the same number of gates (I_{avg}). Ten placement runs were performed using Dragon3.01 and five runs using the greedy method. The p values for blocks $BC[l=1]$ and $ABC[l=1]$ were then calculated from the placement results and then averaged over the results from different placements. In order to calculate the p values, first the average number of absorbed terminals$<T_{abs}[S]>$ is extracted for different orientations of block BC and ABC (shown in Figure 3.3 case $l=1$), where the corresponding block is of size S. The absorbed terminals are those interconnect terminals that are used to connect between the various cells within the block. Then using $<T_{abs}[S]>$, the average number of unabsorbed terminals $<T_{unabs}[S]>$ is calculated for the blocks BC and ABC by using Eq3.12. The unabsorbed terminals refer to those terminals that come out of the block to be connected to the cells outside the block.

$$<T_{unabs}[S]> = S \cdot I_{avg} - <T_{abs}[S]> \qquad (Eq3.12)$$

Eq3.12 is used to get $<T_{unabs}[S]>$ instead of directly counting the terminals of each block because the gates at the periphery (outermost level neighbors) are connected to only one other cell. So if the number of unabsorbed terminals T_{unabs} is counted directly, for blocks that include the peripheral gates, it will be lesser than the actual average values observed in a real netlist. However the number of absorbed terminals T_{abs} will not be affected by the missing terminals of the peripheral gates, because if they do exist, they will have to be connected to the cells outside the block. If they connect to the already existing cells within the block, because these cells share a common parent node, the resulting graph will have a cycle. Because the graph that is constructed is a tree, and if missing terminals do exist, they will not be connected to the already existing cells within the block. The already existing cells share a common parent node in the graph because they are all part of a tree. Therefore, Eq3.12 is used to calculate $<T_{unabs}>[S]$. The p values for blocks BC[l=1] and ABC[l=1] can then be calculated using Eq3.13 with block size S=2 and S=3 respectively, where Eq3.13 is identical to Eq3.4.

$$p[S] = \frac{\log\left(\frac{<Tunabs[S]>}{k}\right)}{\log(S)} \forall S = 2,3 \qquad \text{(Eq3.13)}$$

Figures 3.8 and 3.9 gives the $p[2]$ and $p[3]$ values as a function of the average number of interconnects connected to a cell position extracted from the synthetic graph. The figure also includes the actual values of $p[\text{BC}[l=1]]$ and $p[\text{ABC}[l=1]]$ from the different IBM benchmark circuits. Just as predicted by the theory, the results from the Greedy placement produce the approximate upper bound, while the results from Dragon produce the lower bound values for the *initial exponent pair* $p[\text{BC}[l=1]]$ and $p[\text{ABC}[l=1]]$.
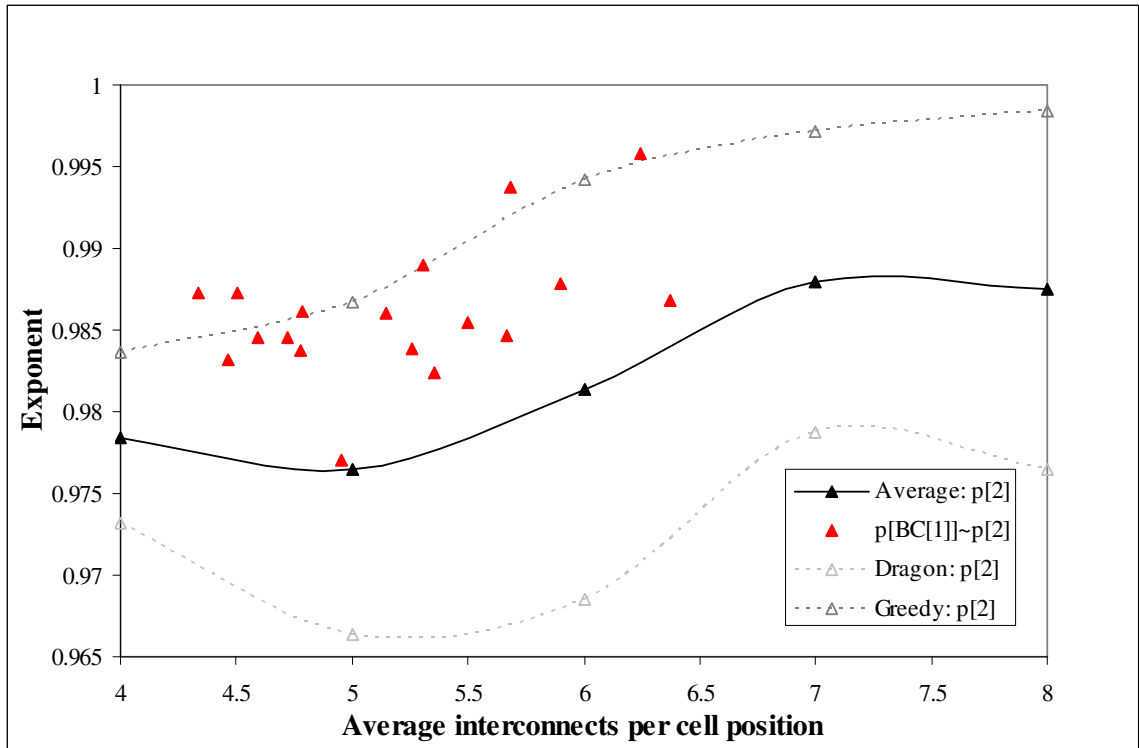
Figure 3.8 Exponent *p*[BC[*l*=1]] vs. average interconnects per cell position
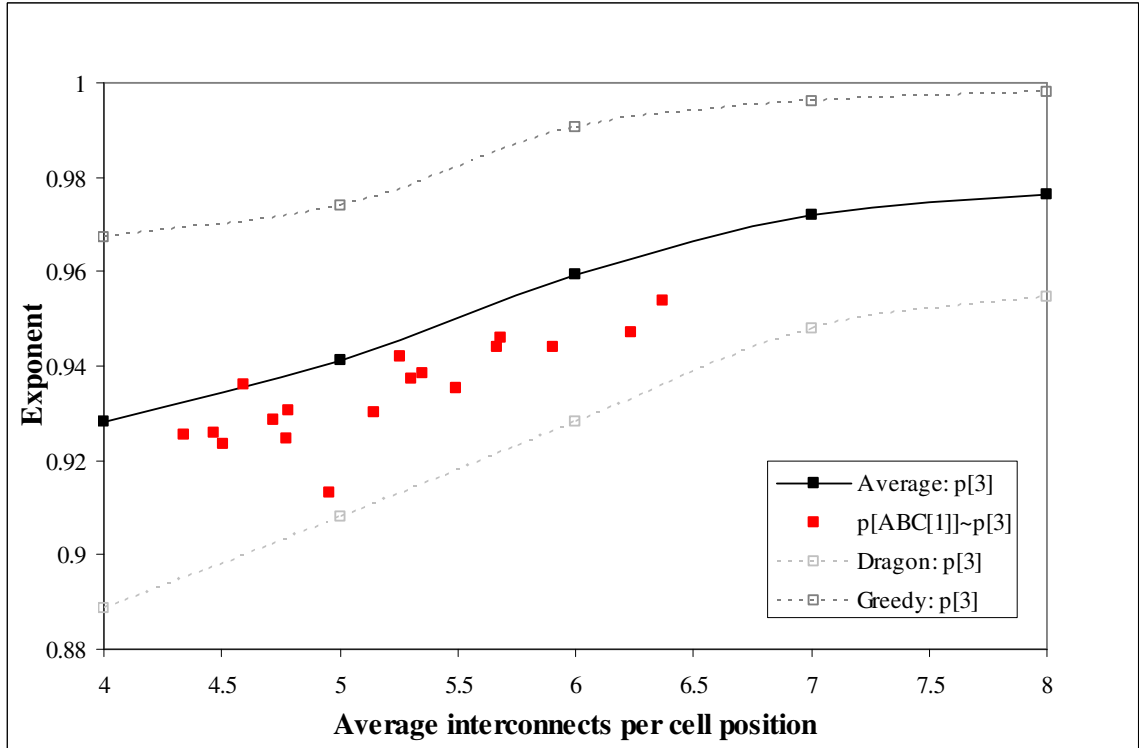
Figure 3.9 Exponent *p*[ABC[*l*=1]] vs. average interconnects per cell position

It can be seen that while the actual $p[BC[l=1]]$ values are closer to greedy results, the $p[ABC[l=1]]$ values are much closer to Dragon results. A possible reason for this could be the fact that the actual netlist is generally not a tree as used in the experiments but a graph with cycles. The presence of cycles could increase the number of short interconnects due to higher connectivity. This in turn could increase the number of terminals absorbed, and consequently the value of $p[ABC[l=1]]$ is reduced. The greedy results of $p[2]$ also exhibit a moderate correlation coefficient of 0.52 against actual $p[BC[l=1]]$ values. A high correlation, with a coefficient of 0.84, was observed between the average results of $p[3]$ and actual $p[ABC[l=1]]$ values. Therefore a lookup table is constructed such that given the average source-sink pair terminal count per gate position, the greedy model results are used to look up the $p[BC[l=1]]$ values and the average of Greedy and Dragon results is used to model $p[ABC[l=1]]$.

Complete p model

The j value to be used in the model for $p[ABC[l]]$ is calculated using Eq3.14 by substituting the values of $p[ABC[l=1]]$ from the initial exponent pair and q into Eq3.10.

$$j = \frac{p[ABC[l=1]]}{\left(N_{ABC[l=1]}\right)^{-q}} \tag{Eq3.14}$$

The term $N_{ABC[l]}$ refers to the size of the block ABC[l] in Eq3.14 and Eq3.15, where Eq3.15 is the p model for block ABC.

$$p[ABC[l]] = j\left(N_{ABC[l]}\right)^{-q} \tag{Eq3.15}$$

It was observed from Figure 3.2 that the blocks BC[l] (same as B[l+1]) exhibit a considerably different p value trend from ABC[l] (same as AB[l+1]). Therefore, the values of $p[BC[l]]$ is estimated by using an approximation technique. From Figure 3.3 it can be seen that there is a BC[l] corresponding to every ABC[l]. It is assumed that each group of $p[BC[1]]$, $p[BC[l]]$ and $p[ABC[l]]$ can be fitted to an expression Eq3.16, which is of the form of Eq3.10.

$$p[Block] = j_{BC[l]}(BlockSize)^{-q_{BC[l]}} \qquad \text{(Eq3.16)}$$

So Eq3.16 is fitted to the pair of $p[ABC[l]]$ and $p[BC[l=1]]$ to yield a $j_{BC[l]}$ (Eq3.20)and $q_{BC[l]}$ (Eq3.21), which is then used to estimate $p[BC[l]]$ (Eq3.21).

$$p[BC[1]] = j_{BC[l]}(N_{BC[1]})^{-q_{BC[l]}} \qquad \text{(Eq3.17)}$$

$$\Rightarrow \log(p[BC[1]]) = \log(j_{BC[l]}) - q_{BC[l]}\log(N_{BC[1]}) \qquad \text{(Eq3.18)}$$

Similarly, $\qquad \log(p[ABC[l]]) = \log(j_{BC[l]}) - q_{BC[l]}\log(N_{ABC[l]}) \qquad \text{(Eq3.19)}$

$$q_{BC[l]} = \frac{\log(p[BC[1]]) - \log(p[ABC[l]])}{\log(N_{ABC[l]}) - \log(N_{BC[1]})} \qquad \text{(Eq3.20)}$$

$$j_{BC[l]} = \frac{p[BC[1]]}{(N_{BC[1]})^{-q_{BC[l]}}} \qquad \text{(Eq3.21)}$$

$$p[BC[l]] = j_{BC[l]}(N_{BC[l]})^{-q_{BC[l]}} \qquad \text{(Eq3.22)}$$

New Wire Length Distribution Model

Given the p values, the number of terminals can be estimated using a modified Rent's expression shown in Eq3.23

$$T[Block] = k(Blocksize)^{p[Block]} \qquad \text{(Eq3.23)}$$

But for a given length, the size of blocks B and C varies based on the position of block A within the layout (see Figure 3.3). This is due to the fact that as the block A gets closer to the periphery of the layout, only a partial part of the blocks B, BC, AB and ABC will lie within the layout. In order to account for these variations in block size, the terminal count of the block of given length is calculated as the weighted average of the terminal counts for various areas of the blocks. The weight used is the frequency $F[N_{Block[l]}]$ of each possible area ($N_{Block[l]}$), for a given block ($Block[l]$) of a given length (l), which is counted using a simple counting algorithm. The algorithm counts the frequency of the block sizes for each of the blocks BC[l] and ABC[l] and for each length l, based on all the possible positions of block A.

$$\text{For each length } l: \quad T_{avg}\big[Block[l]\big] = \frac{\displaystyle\sum_{\forall N_{Block[l]}}\big(T[Block[l]]\cdot F[N_{Block[l]}]\big)}{\displaystyle\sum_{\forall N_{Block[l]}} F[N_{Block[l]}]} \qquad \text{(Eq3.24)}$$

The results from Eq 3.24 can then be used along with law of conservation of terminals to calculate the number of interconnects between each pair of block A and Block C separated by a distance l as shown in Eq3.25 or Eq3.26. The interconnect length distribution is then calculated as a product of the total number of cell positions $N_{cellpositions}$ and $I_{A\text{-}to\text{-}C}[l]$ as shown in Eq3.27, where the total cell positions also gives the total number of possible positions of block A, and hence the total number of pairs of block A and block C.

$$I_{A-toC}[l] = \big(T_{avg}\big[BC[l]\big] + T_{avg}\big[AB[l]\big] - T_{avg}\big[ABC[l]\big] - T_{avg}\big[B[l]\big]\big) \qquad \text{(Eq3.25)}$$

$$I_{A-toC}[l] = \big(T_{avg}\big[BC[l]\big] + T_{avg}\big[ABC[l-1]\big] - T_{avg}\big[ABC[l]\big] - T_{avg}\big[BC[l-1]\big]\big) \qquad \text{(Eq3.26)}$$

$$I[l] = I_{A-to-C}[l]\cdot N_{cellpositions} \qquad \text{(Eq3.27)}$$

Variability model

It was seen in the previous chapter that there exists a variation in the distribution of the interconnect lengths from run-to-run and from placement tool-to-tool. The coefficient of variation ($CoV(I[l])$) of the distribution was also observed to have relation with the distribution. The plot highlighting the relation between the two is shown in Figure 3.10, and an expression for the coefficient of variation as a function of the interconnect length distribution is derived from this relation is given by Eq3.29.

$$
\begin{aligned}
&\text{If } (I[l] > 1000) \quad \log(CoV(I[l])) = 0.1057\cdot\log(I[l]) - 1.6104 \\
&\quad\text{else} \quad \log(CoV(I[l])) = 0.0401 - 0.4398\cdot\log(I[l])
\end{aligned}
\qquad \text{(Eq3.28)}
$$

$$
\begin{aligned}
&\text{If } I[l] > 1000 \\
&\qquad CoV(I[l]) = \frac{I[l]^{0.1057}}{10^{1.6104}} \\
&\text{else} \\
&\qquad CoV(I[l]) = 10^{0.0401}\cdot I[l]^{-0.4398}
\end{aligned}
\qquad \text{(Eq3.29)}
$$
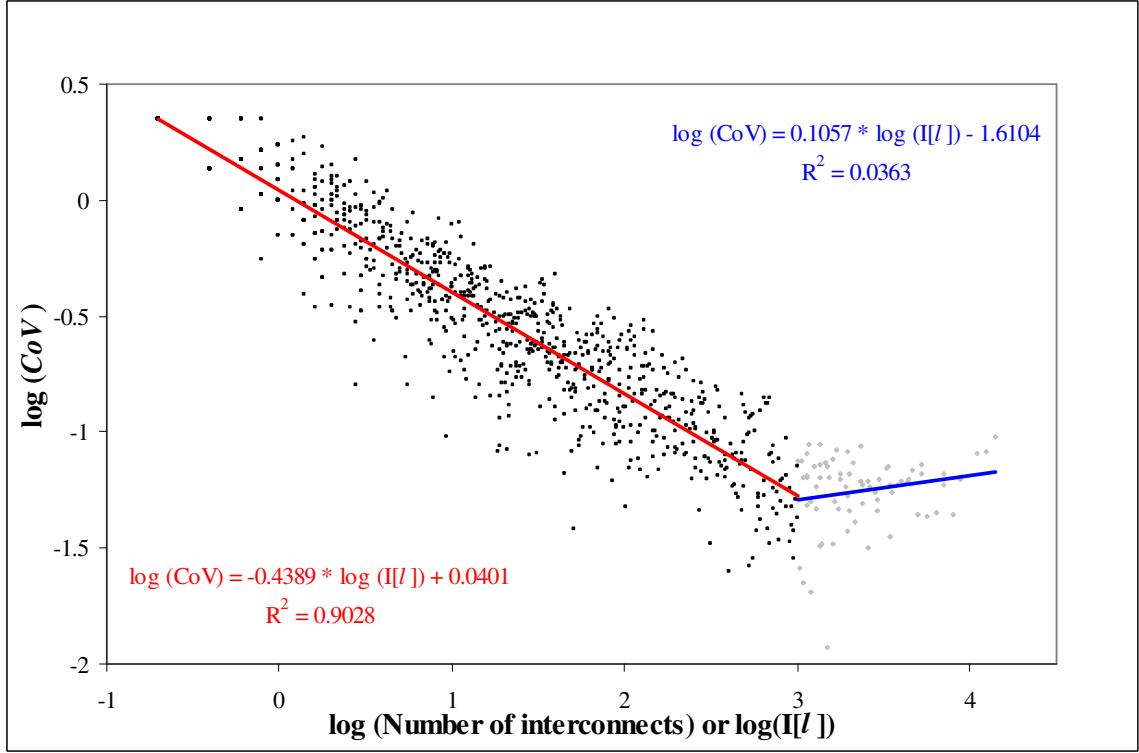
Figure 3.10 Coefficient of variation as a function interconnect length distribution

Given the expression for the coefficient of variation CoV(I[l]), the standard deviation of the interconnect length distribution σ(I[l]) can be calculated using Eq3.30.

$$\sigma(I[l]) = I[l] \cdot CoV(I[l]) \qquad \text{(Eq3.30)}$$

Using the expression for standard deviation and the calculated I[l] values, an approximate upper and lower bound can be calculated for the length distributions. For each length l, the upper bound is estimated as three standard deviations above the calculated I[l] value and the lower bound is estimated as three standard deviation below the I[l] value.

Interconnect length distribution calculation methodology

The procedure for calculating the distribution for a given netlist is given below:

Step 1.   Get number of interconnects I

Step 2.   Get number of cell positions in the layout N$_{cellposition}$

Step 3.   Calculate k using Eq3.3

69

Step 4.   *Calculate q from Eq3.11*

Step 5.   *Get the initial exponent pair p[ABC[1]] and p[BC[1]] using k in the look up table model*

Step 6.   *Calculate j using Eq3.14*

Step 7.   *For each length l (1 to X+Y-2) (X, Y dimensions of the layout grid)*

        a. Calculate p[ABC[l]] using Eq3.15

        b. Calculate T[ABC[l]] using Eq3.23

        c. If(l=1)

           i.   p[BC[l]] = p[BC[1]]

          else if (l>1)

           i.   Calculate $q_{BC[l]}$ using Eq3.20

           ii.   Calculate $j_{BC[l]}$ using Eq3.21

           iii.   Calculate p[BC[l]] by interpolation using Eq3.22

        d. Calculate T[BC[l]] using Eq3.23

        e. Calculate the frequency of the areas of block BC[l] $F[N_{BC[l]}]$

        f. Calculate the frequency of the areas of block ABC[l] $F[N_{ABC[l]}]$

        g. Calculate the average terminal count T_{avg} [BC[l]] using Eq3.24

        h. Calculate the average terminal counts T_{avg}[ABC[l]] using Eq3.24

        i. Calculate $I_{A\text{-}to\text{-}C}$[l] using Eq3.26

        j. Calculate I[l] using Eq3.27

Step 8.   *Renormalize distribution to match the total number of interconnects if necessary*

Step 9.   *Calculate CoV (I[l]) usingEq3.29 and σ(I[l]) using Eq3.30*

Step 10.   *Calculate the approximate bounds to I[l] as I[l]+ 3σ(I[l]) and I[l]- 3σ(I[l])*

It should be remembered that the parameter *q* in (Step 4) is derived from a model developed based on values extracted from a set of sample benchmark results.

The interconnect length distribution result from the new model (New) is shown in Figure 3.11 for IBM10 benchmark circuit along with the results from the DDM model. The figure also includes the results from the simulated annealing based placement (SA). It can be clearly seen that the new model predicts more accurately for the shorter and medium length interconnects. But as the length increases the prediction accuracy of the new model suffers.
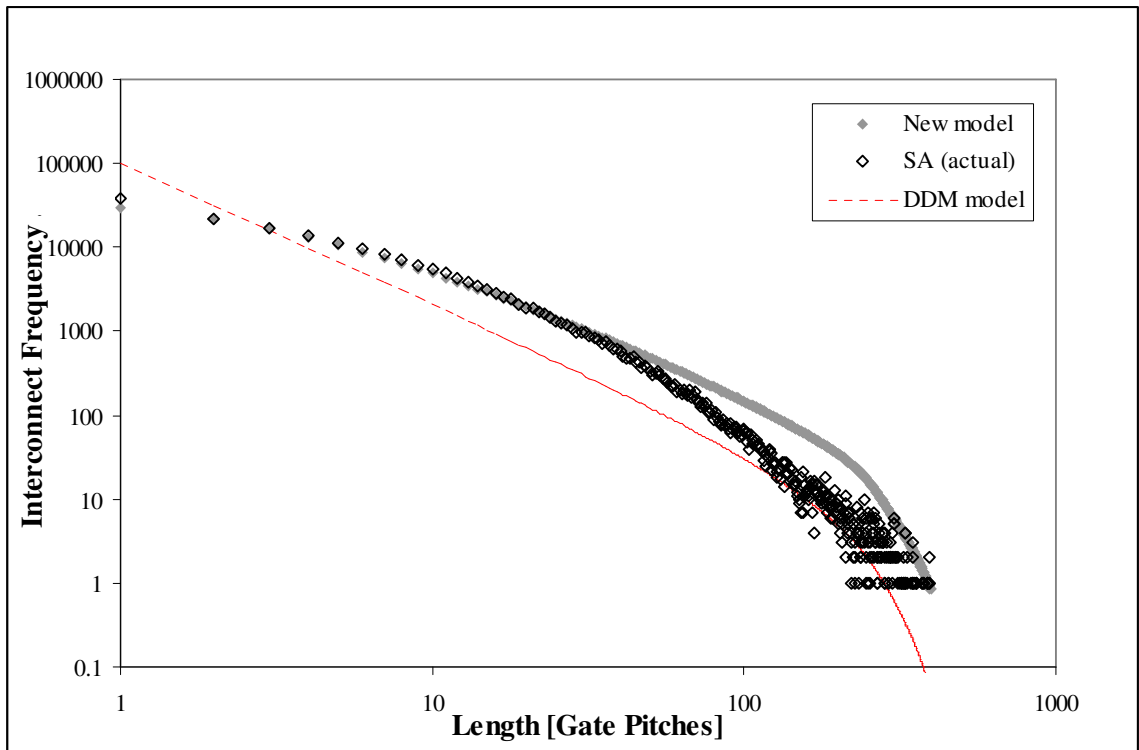
Figure 3.11 New model results against the DDM model results for IBM10

New model vs. DDM model

A closer evaluation of the improvement provided by the new model can be made using the analysis of the DDM model from Chapter 2. The DDM model was analyzed by dividing the distribution into three zones, and estimating its zone attributes. It was observed that while only slightly more than one third of the interconnects actually belong to zone 1, comprised of very short interconnects, the DDM model overestimates this to

71

almost double its original frequency. Consequently due to normalization, the nearly two thirds of the interconnects present in zone 2 was found to be under estimated to approximately half its original value and the less than 0.5% of interconnects in zone 3 was found to be underestimated by a factor of 35.14% respectively. These zone attribute values are given again in Table 3.1 along with the zone attribute values, of these same zones as defined by the DDM model, calculated with the new model. It can be seen that with the new model, the absolute error in zone 1 and 2, which comprises of more than 99.5% of interconnects, is drastically reduced to 9.47% and 5.44% on an average from +94% and -53% respectively. This corresponds to a 10x reduction in error in interconnect length distribution prediction of these zones.

However, with the new model, for zone 3 comprising less than .5% of interconnects the error is increased to 426%. This is due to the error in modeling the p[BC[l]] values for large block sizes. It was observed that as the block sizes become very large, the difference between the number of terminals of a block BC and block ABC tends to zero. In other words, when the block A with a single cell at the center of the semi-manhattan circle is added to the block BC that is the semi manhattan circle, on an average, half of the single cell's terminals in block A is connected to the cells within the block BC and the remaining half to the cells outside. Consequently the number of terminals of block BC and ABC tend to be the same. But the approximated interpolation based model for p[BC[l]] used in the new model is not a very good representation of the actual p[BC[l]]values. This is the primary cause for the huge error among the distribution of interconnects of very long lengths. A secondary issue is that for the distribution of interconnects vary widely for very long interconnects. For example, there will be no interconnects of certain length and there will be few for the immediately adjacent length. But the model always predicts this value to be positive, which again can cause the error to be high for zone 3.

TABLE 3.1
LENGTH DISTRIBUTION'S ZONE ATTRIBUTES DDM VS. NEW MODEL

Zone	Circuit	EndL	%IC*	%ErrorCD*	
			Actual	DDM	New Model
	IBM01	2	39.42	+91.77	-22.54
	IBM02	3	31.74	+113.48	+3.83
1	IBM03	3	41.99	+71.79	-5.73
	IBM04	2	34.20	+85.88	-8.33
	IBM05	4	33.65	+109.07	+6.92
	Average	2.8	**36.21**	**+94.39**	**9.47(absolute)**
	IBM01	66	60.33	-59.86	+10.73
	IBM02	74	67.04	-53.39	-7.22
2	IBM03	118	57.67	-52.01	+2.68
	IBM04	128	65.56	-44.72	+2.88
	IBM05	203	66.26	-55.33	+3.70
	Average	117.8	**63.37**	**-53.06**	**5.44(absolute)**
	IBM01	218	0.25	-25.68	+987.73
	IBM02	275	1.20	-19.22	+300.28
3	IBM03	295	0.32	-45.85	+262.37
	IBM04	323	0.23	-22.20	+415.02
	IBM05	335	0.07	-63.66	+166.25
	Average	289.4	**0.42**	**-35.14**	**+426.33**

* %IC: % Interconnect count (cumulative distribution) in zone; %ErrorCD: % Error in cumulative
distribution between model results and actual simulated annealing based placement results

New model with variability

Figure 3.12 shows the results produced by the new model with the variability model incorporated into it for IBM18 benchmark circuit. It gives the bounds based on both three times the standard deviation and two times the standard deviation. The figure also gives the results from the simulated annealing based placement result. It can be seen that with the variability model included, the actual distribution of most of the lengths fall within the bounds. Similar trend was observed in the placement results of a lot of other benchmark circuits. However, even with the variability model incorporated into the new model, sometimes for some placement tools, the distributions are still difficult to accurately predict. This can be seen from Figure 3.13.
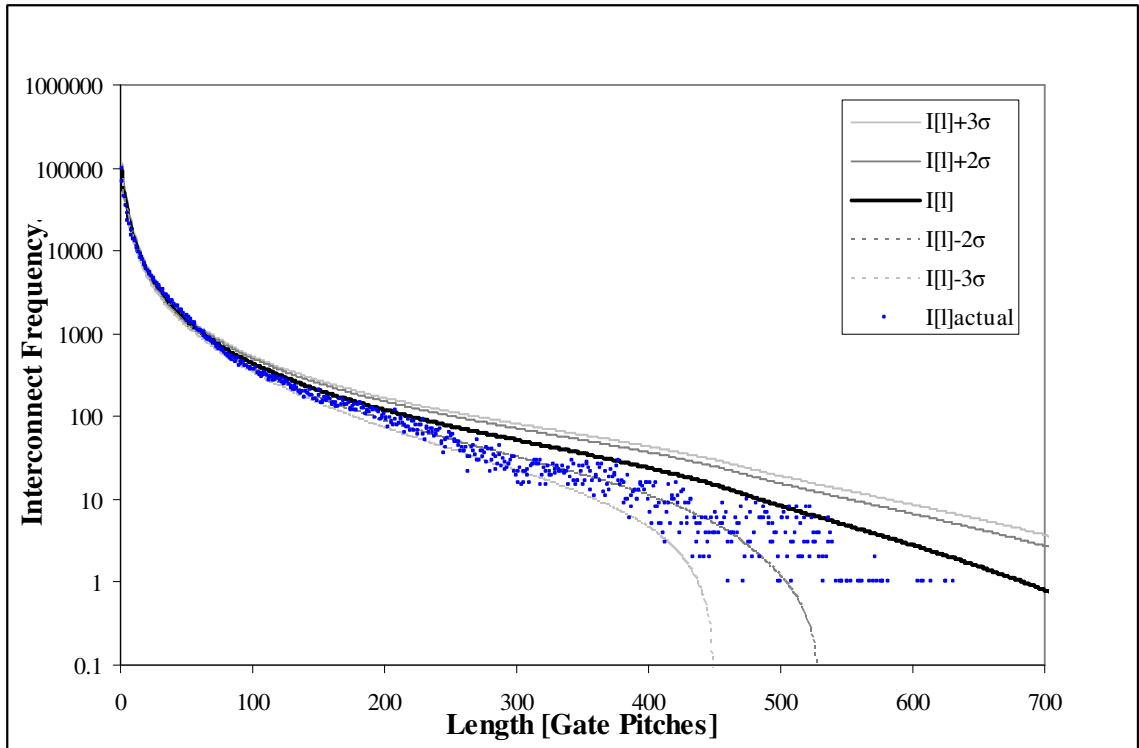
Figure 3.12 Results of new model with variability for IBM18
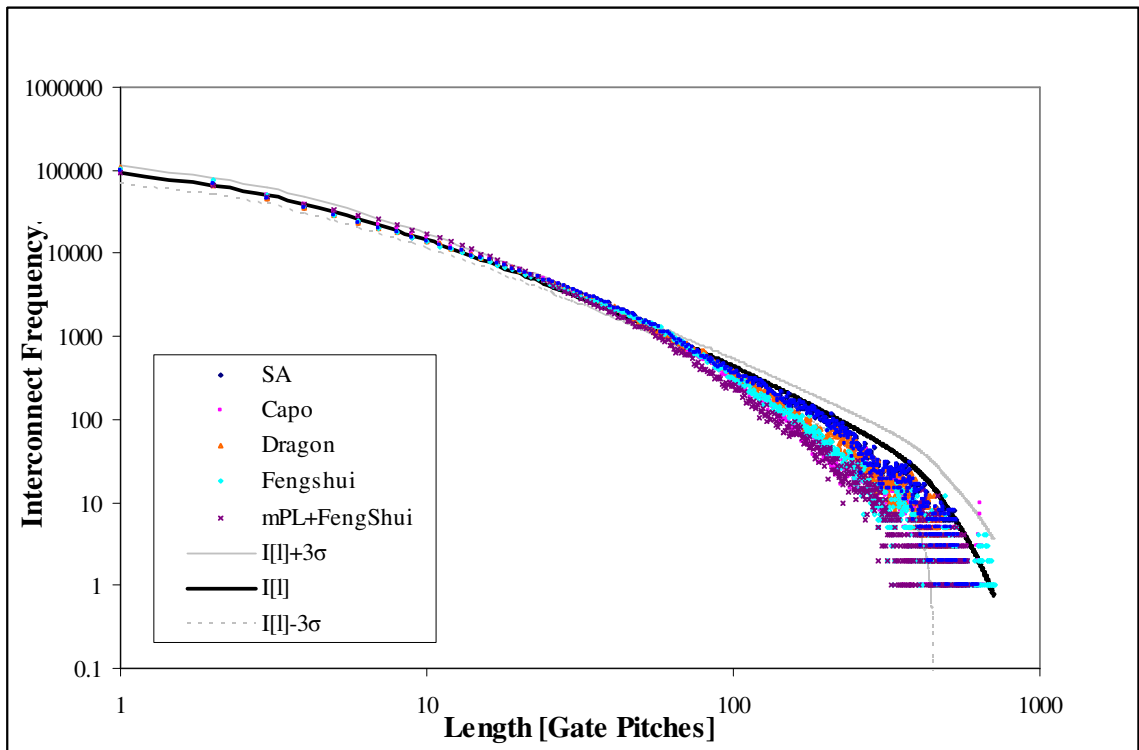
Figure 3.13 New model against results from different placements for IBM18

Figure 3.13 gives the distribution plot of the result from the new model compared against the results from five different placement tools for IBM18 benchmark circuit. It can be seen from this figure that, for the distributions calculated from the results of some of the placement tools, the actual distributions fall outside the variability range for very long interconnects. This proves the earlier stated theory about over prediction for very long interconnects due to error in modeling $p[BC[l]]$. Therefore, it is possible to improve the accuracy further by improving the model for predicting $p[BC[l]]$.

Summary

Using the limitations to prediction accuracy identified from earlier investigations, a new interconnect length distribution model is developed as an improvement over the DDM model. The key corner stones of the new model are (i) a new model for parameters $p[ABC[l]]$ and $p[BC[l]]$, which are used in a modified Rent's rule to calculate the Interconnect-terminal-to-Block-size relation of a placed netlist, (ii) a look-up table that provides the p values of the two smallest block sizes used in model ($p[ABC[l=1]]$ and $p[BC[l=1]]$) and (iii) incorporation of the empirical variability model into the wire length distribution model. The result is a highly accurate prediction of the distributions of the shorter wires. However the model still suffers from a loss of accuracy for the longer wires.

CHAPTER 4

LIMITATIONS OF MICROSCOPIC PREDICTION

Early and accurate individual wire length prediction (i.e. microscopic prediction) could solve many of the problems due to the lack of physical design information at early stages of the design cycle. This chapter deals with the investigations pertaining to the limitations of predicting the length of the wires individually. These investigations are carried out from two different perspectives. First, the set of investigations discussed in the initial part of the chapter pertain to the difficulties in microscopic length prediction independent of the prediction methodology. The second part of the chapter reviews the state-of-the-art in current microscopic wire length prediction models to develop an understanding of limitations inherent in current prediction methods.

Microscopic Repeatability

To be able to predict the length of a wire, the wire must have consistently similar lengths in the optimized placement solutions obtained using different placement tools and at the end of different placement runs. If the length of the wire varied widely from one placement solution to another, prediction of the length would be next to impossible. Therefore, it is imperative to find out if there are wires that have consistently similar lengths or in other words *repeatable wires* from one placement solution to another.

If there are such wires with consistently similar lengths, answers to the following questions will shed more light on the limitations for individual wire length prediction: (i) what fraction of wires have similar lengths? (ii) what are the net degrees of these wires? (iii) what is the length distribution of these wires? This will also enable the development of more robust prediction models based on a solid understanding of the possibilities for

microscopic length prediction. Therefore, in the following subsections, results from the investigations that were carried out to evaluate these attributes are explained.

Since there are three different wire models: net, interconnect and connection, the investigations are performed for each of the three wire models. The benchmarks placed using the five different placement tools viz., In-house Simulated Annealing, Dragon3.01, Capo8.8, FengShui5.0 and finally mPL5.0 based global placement combined with FengShui5.0 based detailed placement are used for these investigations.

Investigation methodology

The following procedure is used to analyze the repeatability of wire lengths of all the three wire models (i.e. net, interconnect and connection). For each netlist, the wire length of each wire is extracted from each of the five different placement solutions. The semi-perimeter bounding box length is extracted for a net, and the length between the two terminals of the interconnect or the connection is extracted for an interconnect or a connection wire model. Let $L_{act, w, placement}$ be the length of the wire w obtained from a given placement solution. Then, the average length of each wire $L_{avg, w}$ over the five placement solutions is calculated as

$$L_{avg, w} = \frac{\sum\limits_{\forall \, placements} L_{act, w, placement}}{5} \qquad \text{(Eq4.1)}$$

The maximum length deviation Δ_w for the given wire w is calculated as the maximum of the absolute difference between the actual length of the wire from each of the placement solution and its average length.

$$\Delta_w = \max \left| L_{act, w, placement} - L_{avg, w} \right| : \forall placements \qquad \text{(Eq4.2)}$$

Intuitively, it could be stated that the wire w has a length that is highly repeatable, and therefore predictable, if maximum length deviation for that wire is very small. But there is no clear guideline as to what constitutes a small enough margin for the wire to be considered predictable. A margin parameter Δ_m is therefore introduced and used as a

guideline to define the margin, which has to be met by the maximum length deviation of a wire to be considered as predictable. This margin parameter is varied as a fraction δ of the maximum length of all the wires in a given netlist as shown in Eq4.3.

$$\Delta_m = \delta \cdot \left(\max | Lact, w, p | \forall w, \forall p \right): \forall 0.01 \leq \delta \leq 0.2 \qquad \text{(Eq4.3)}$$

Given a margin parameter, the wires that have maximum length deviation Δ_w less than the margin parameter Δ_m can be identified, and these wires represent the predictable wires. Once these wires are identified, the various statistics such as the percent of wires that are predictable, their length distribution and the degrees of the nets from which these wires are obtained can be extracted. These results are then used to understand the limitations to microscopic prediction for each of the three wire models are discussed in the following subsections.

Predictability of nets

The variation of the percentage of nets that are classified as repeatable for various values of the fraction δ is shown in Figure 4.1. As the fraction δ is increased the margin parameter is increased and consequently the number of nets with absolute maximum deviation less than the margin increases steeply at first and then slowly. It can be seen from this figure that for a very small value of 0.01 for the fraction δ approximately 40% of the nets have semi perimeter bounding box net lengths that are highly repeatable.

For the sake of simplicity, it is assumed that this fraction of 0.01 corresponds to the margin of allowable repeatability. Based on the results from the margin set by this fraction value, the percent share of repeatable nets as a function of their length and their net degrees are extracted, along with the percentage of nets that are repeatable for each net degree It should also be noted that this fraction value of 0.01 corresponds to a very narrow margin of $\pm$ 1 to 3 gate pitches, for the five benchmarks analyzed based on the maximum net length in the netlist.
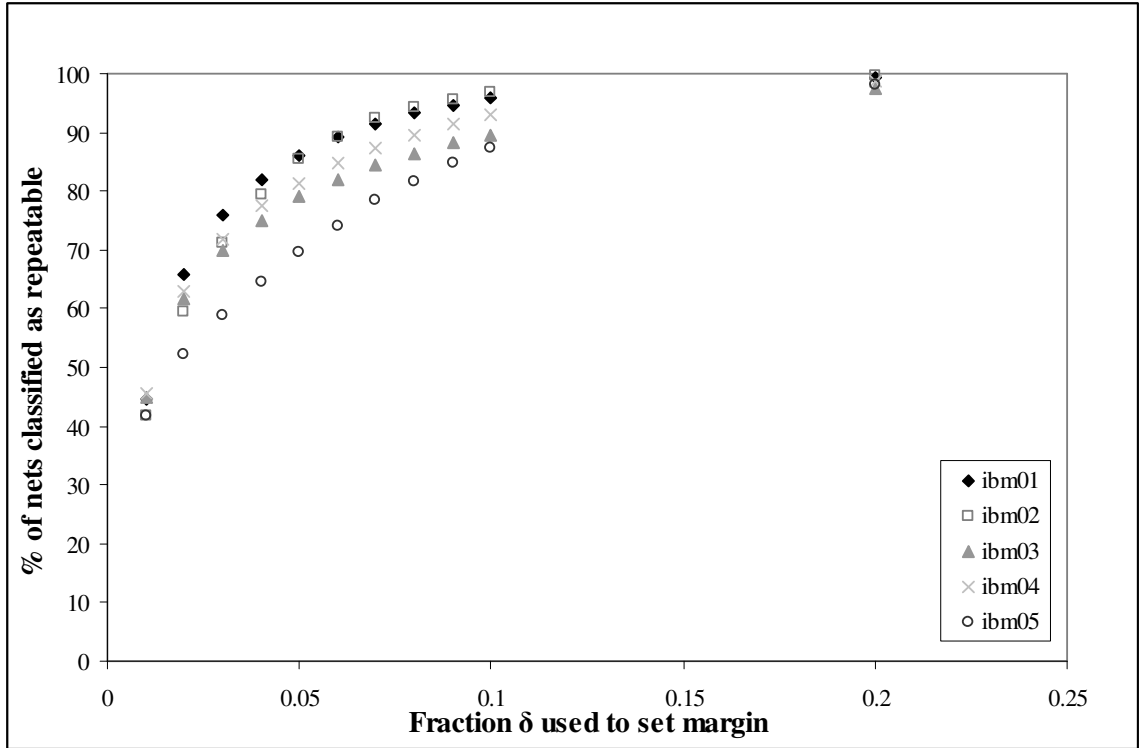
Figure 4.1 Percent of repeatable nets vs. fraction δ

Figure 4.2 depicts the cumulative percent share of repeatable nets as a function of degree of the net for a fraction $\delta = 0.01$. It can be seen from this figure that more than 90% of the highly repeatable 40% of the nets are from nets of degree 2 or 3. Figure 4.3, on the other hand gives the percent of nets classified as repeatable for each net degree for the same fraction value of $\delta = 0.01$. It can be observed that in all the test cases nearly 60 to 70% of the 2 pin nets are highly repeatable, while only a much lesser percent of the higher degree nets are repeatable. Finally, Figure 4.4 plots the relation between cumulative percent of repeatable nets and the average length of the net for the fixed fraction value of $\delta = 0.01$. It can be observed from this figure that nearly all, roughly 90%, of the repeatable 40% of the nets are very short and are of length less than 5 gate pitches. *Based on these results, it can be summarized that slightly less than half the nets are predictable and almost all of them are from shorter nets and have smaller net degrees.*
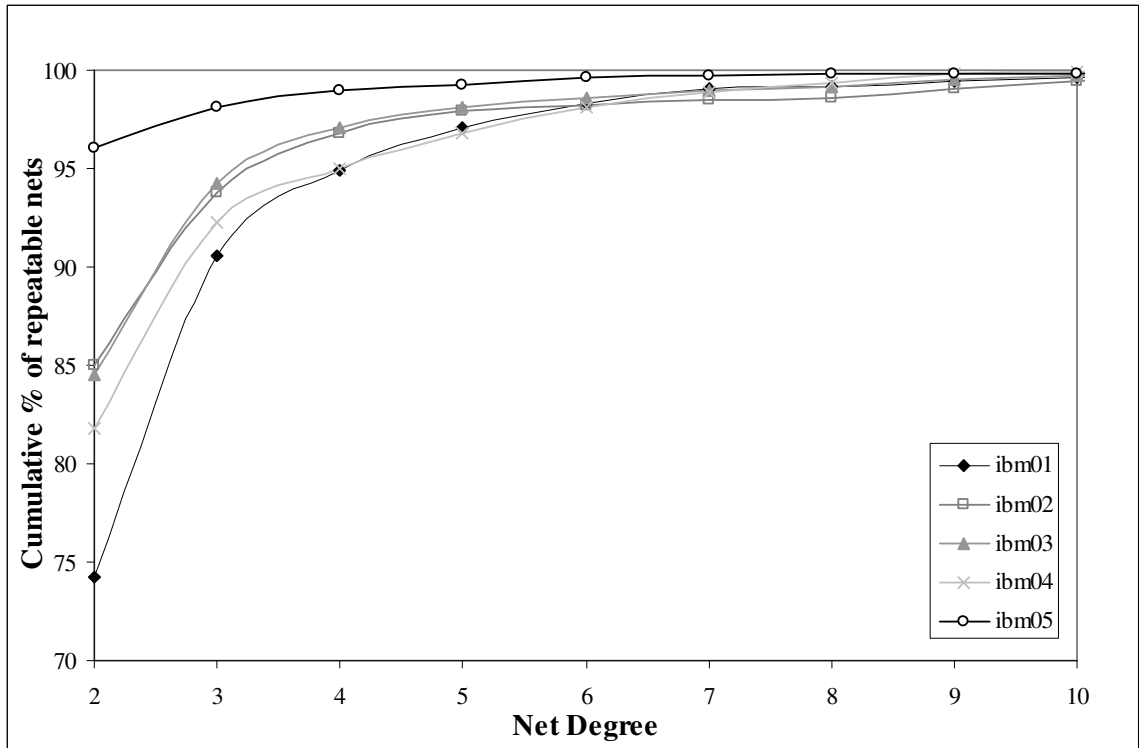
Figure 4.2 Cumulative percent share of repeatable nets vs. net degree
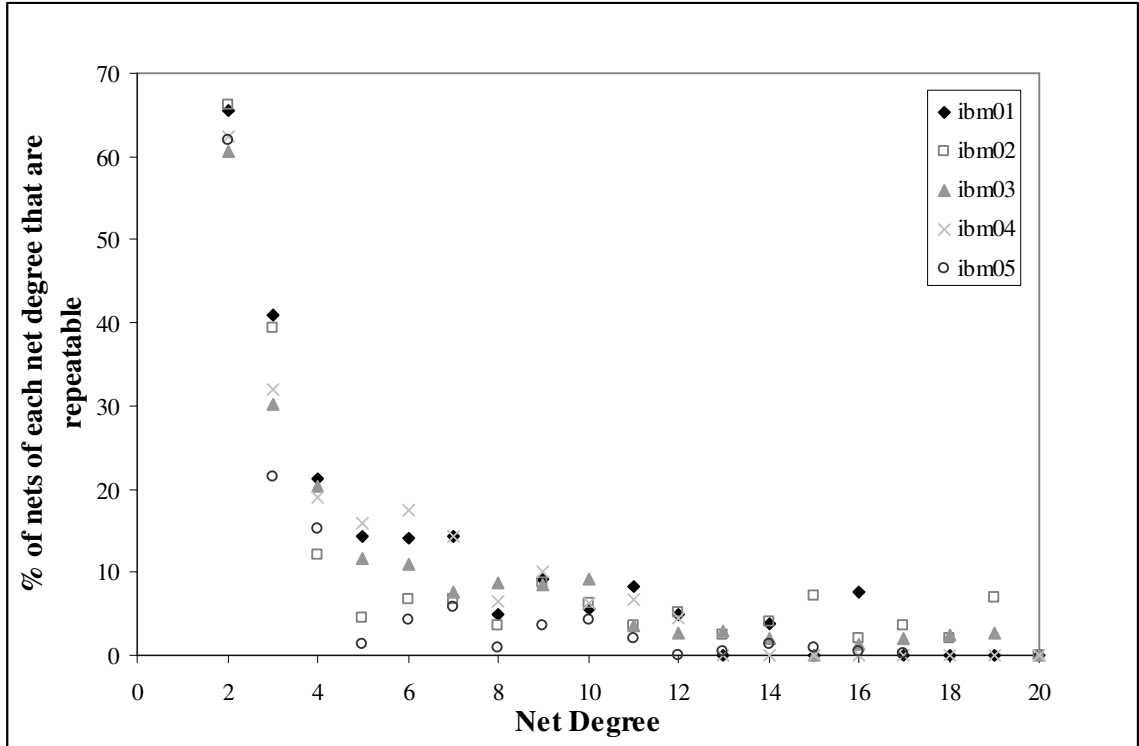
Figure 4.3 Percent of repeatable nets among each net degree for $\delta = 0.01$
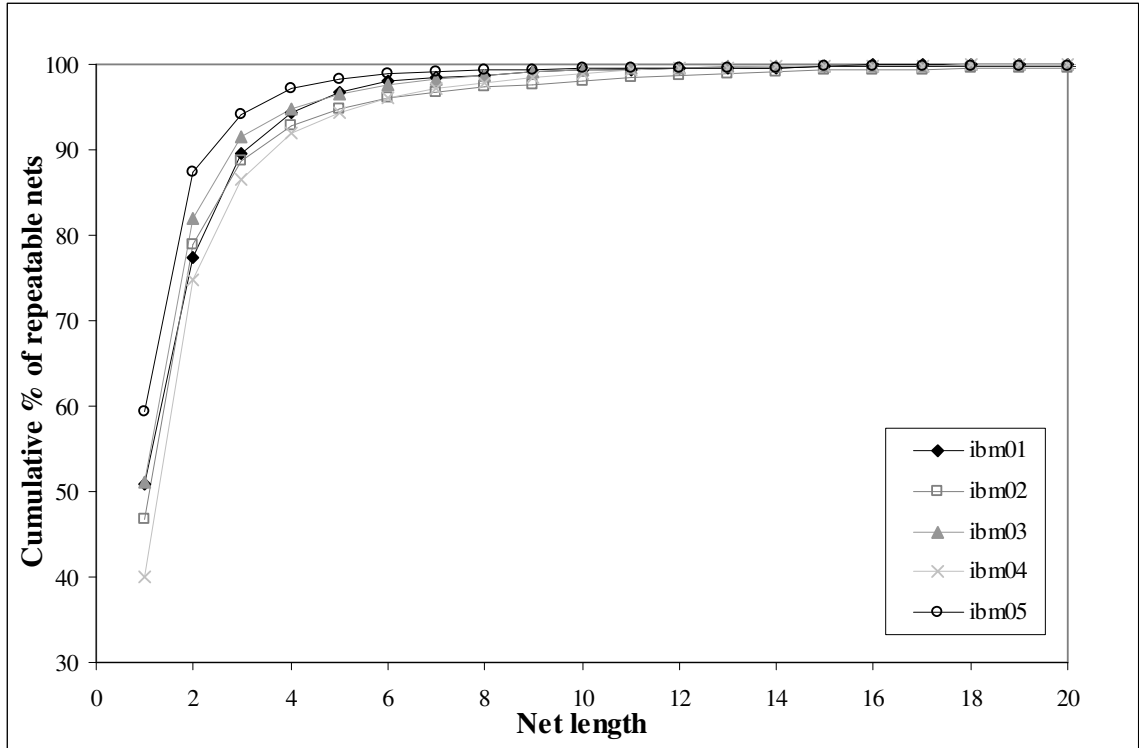
Figure 4.4 Cumulative percent share of repeatable nets as a function of net length

Predictability of interconnects

The variation of percent of interconnects that are classified as repeatable for various values of the fraction δ is shown in Figure 4.5. Similar to the relation observed in the nets, as the fraction δ is increased the number of interconnects with absolute maximum deviation less than the margin increases steeply at first and then slowly. It can be seen from this figure that approximately 20 to 35% of the interconnects have lengths that are highly repeatable. This corresponds to a very small δ value of 0.01, which in turn corresponds to a very narrow margin of ± 1 to 3 gate pitches from the average length. Figure 4.6 shows the cumulative percent share of these highly repeatable interconnects as a function of degree of the net of which the interconnect is part of. It can be seen from this figure that unlike nets, a lesser fraction of approximately 55% to 65% of the highly repeatable interconnects are from nets of degree 2 or 3.
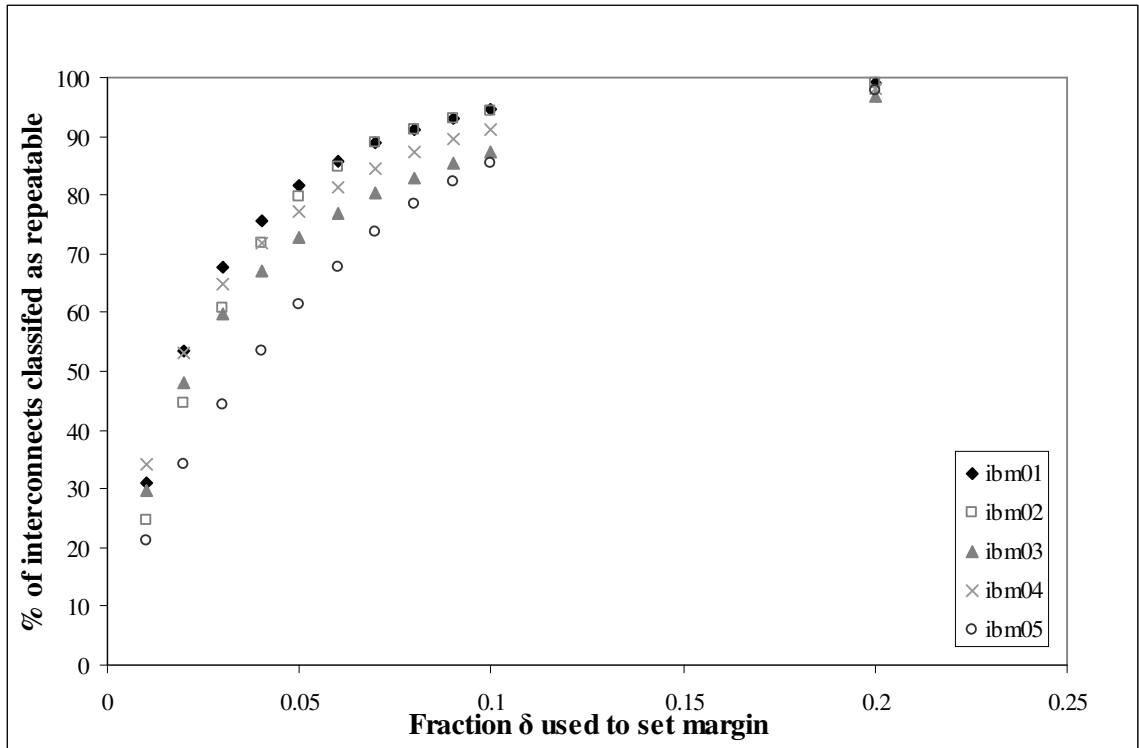
Figure 4.5 Percent of repeatable interconnects vs. fraction δ
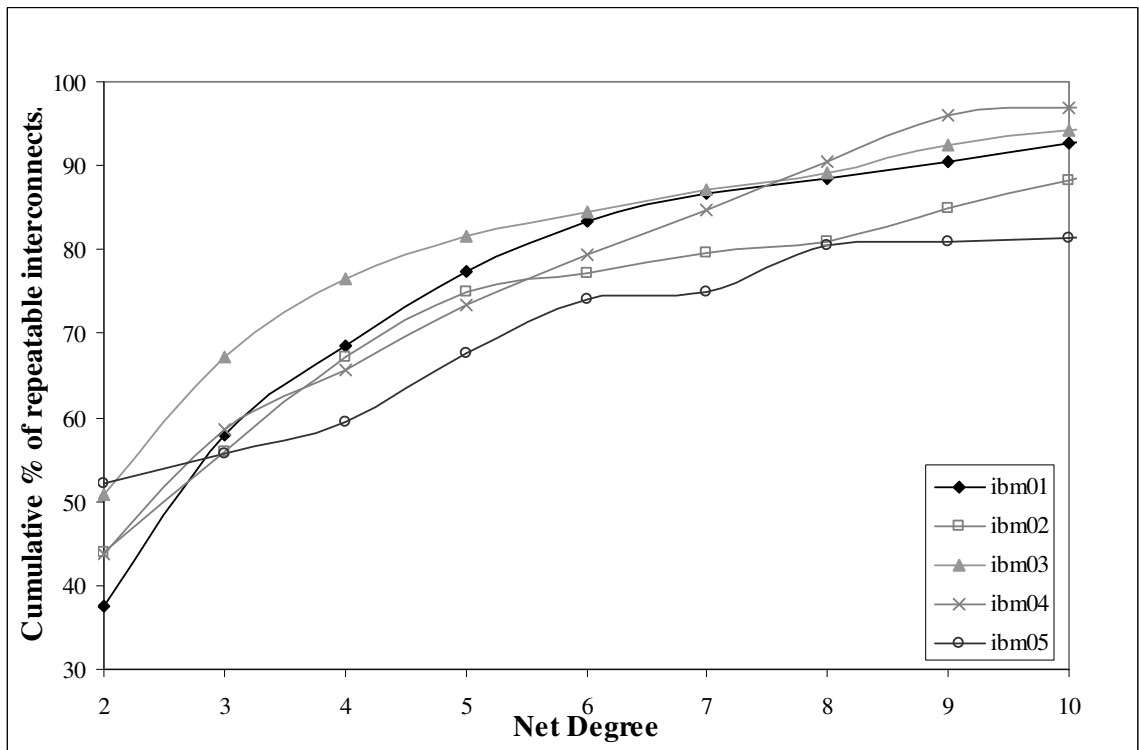
Figure 4.6 Cumulative percent share of repeatable interconnects vs. net degree

Finally, Figure 4.7 gives the cumulative percent of repeatable interconnects as function of interconnect length. Similar to the relation observed for nets, this figure shows that approximately 90% of the repeatable interconnects are of length less than 5 gate pitches. However, it should be noted that the percent of interconnects that are highly repeatable is less than the percent of repeatable nets.
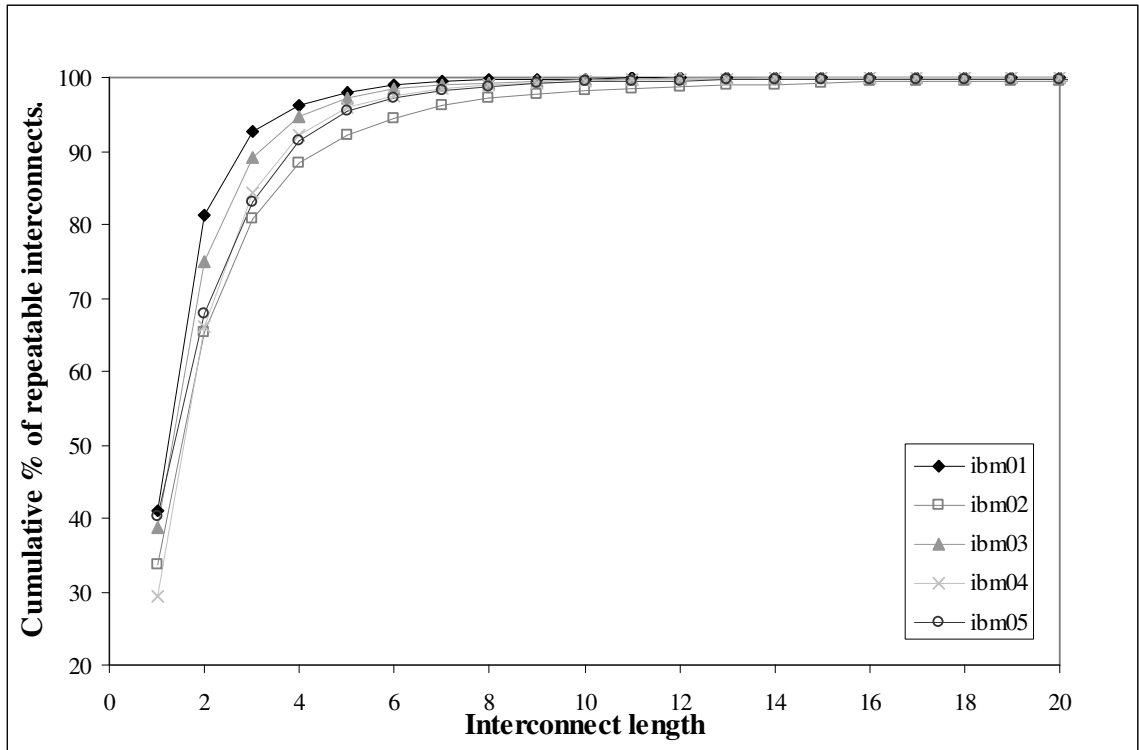
Figure 4.7 Cumulative percent share of repeatable interconnects vs. length

Predictability of connections

The variation in percent connections that are classified as repeatable for various values of the fraction δ is shown in Figure 4.8. Similar to the relation observed in the nets and interconnects, as the fraction δ is increased, the number of connections with absolute maximum deviation less than the margin increases steeply at first and then slowly. Similar to nets and interconnects, the highly repeatable connections are identified with the δ value set at 0.01. It can be seen that only 10 to 25% of the connections have lengths that are highly repeatable, which is smaller than that observed for interconnects and nets.
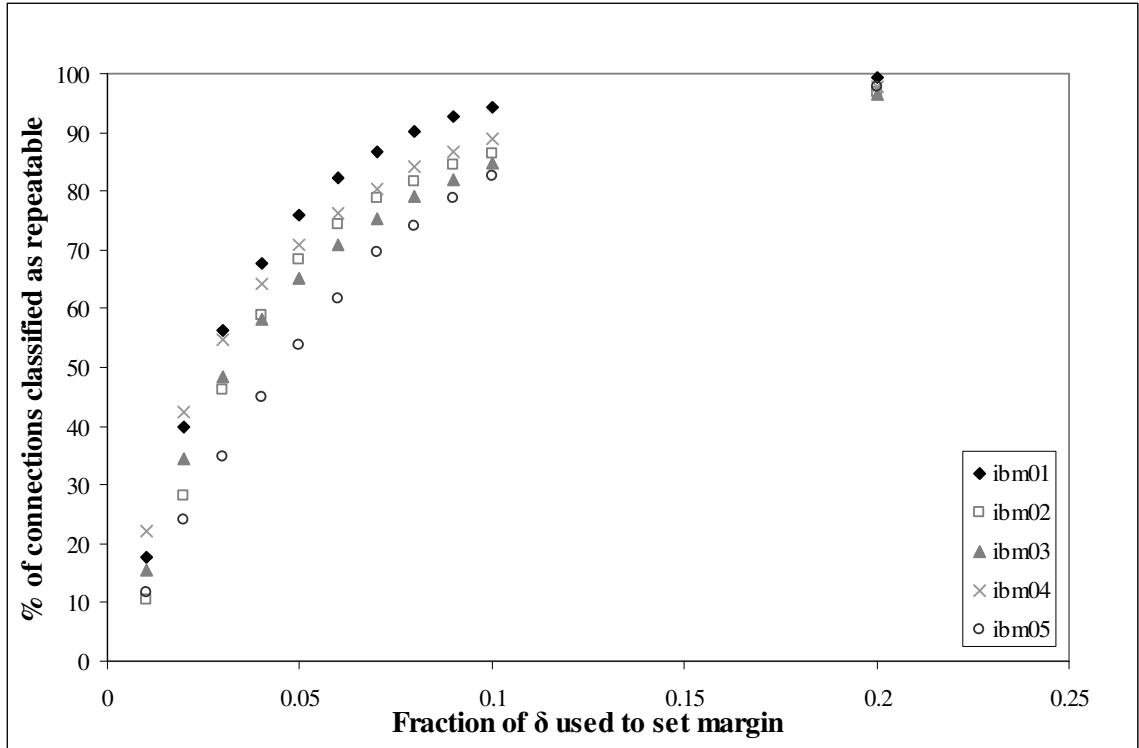
Figure 4.8 Percent of repeatable connections vs. fraction δ

Figure 4.9 shows the relation between the cumulative percent share of the highly repeatable connections and the degree of the net of which the connection is part of. It can be seen from this figure that 20% to 35% of the repeatable connections are from nets of degree 2 or 3. Finally, Figure 4.10 gives the cumulative percent share of connections that are classified as repeatable as a function of its length, and it is observed that at least 70% of the connections classified as highly repeatable are of length less than 5 gate pitches.

A careful analysis of all of the above data revealed that, on an average, there are at least 28% more interconnects that are repeatable than the sum total of the interconnects of the repeatable nets. Similarly, on an average, at least 51% more connections are repeatable than the sum total of the number of connections of the repeatable nets. Bosed on this data it can be inferred that although nets have a higher prediction probability of approximately 40%, more information about the distances between the various terminal gates of the wires can be gleaned by trying to predict connections than nets.
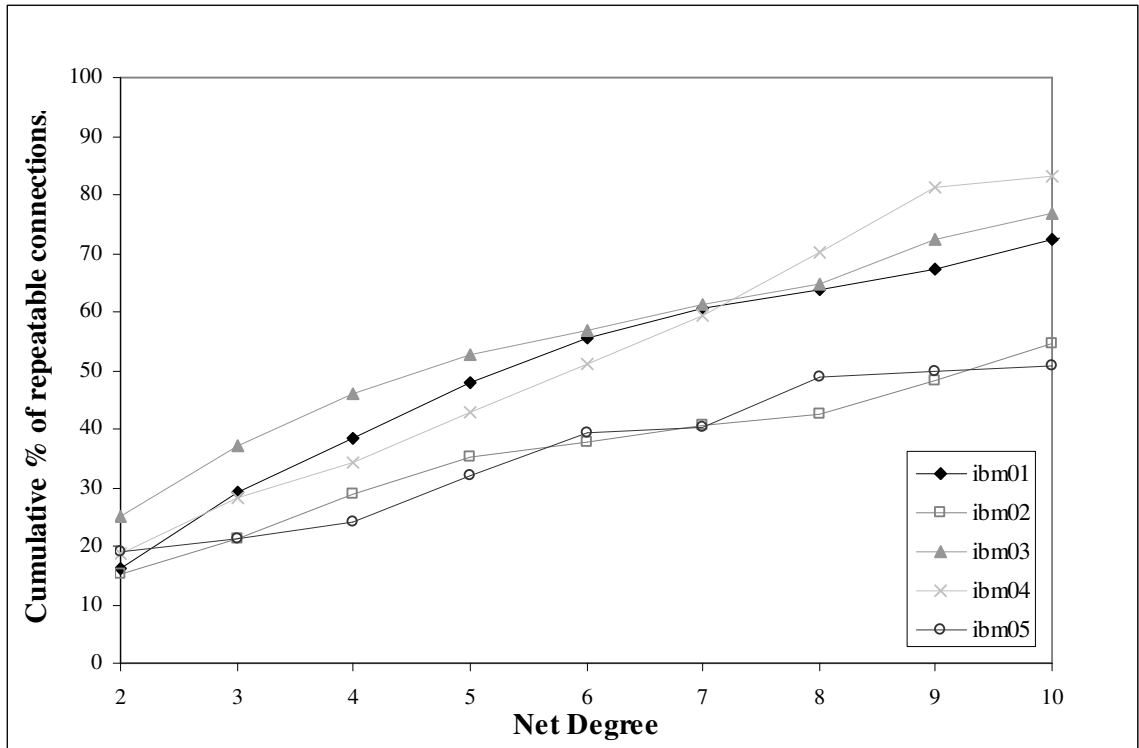
Figure 4.9 Cumulative percent share of repeatable connections vs. net degree
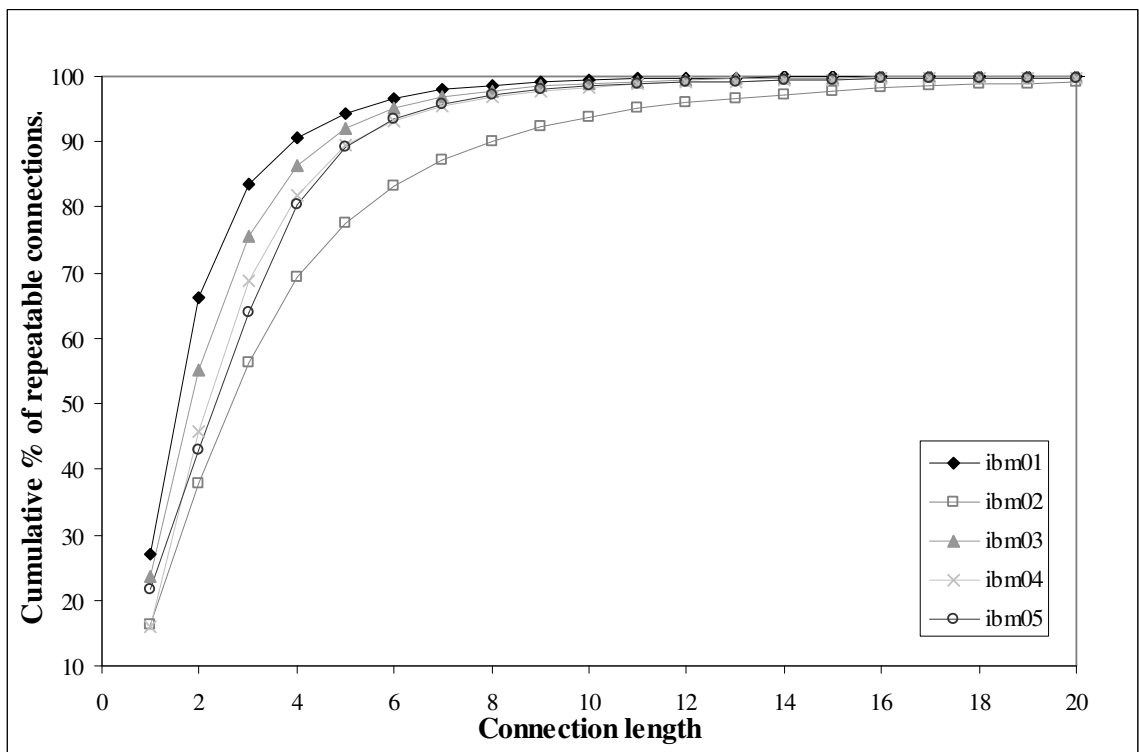
Figure 4.10 Cumulative percent share of repeatable connections vs. length

Based on the above results it is possible to conclude that *microscopic prediction of wire length may be inherently difficult to achieve for all of the wires in a netlist.* However, there certainly remains a subset of the wires in netlist that are easier to predict, and most of them are very short in nature and are part of nets of smaller degrees. Finally, although a larger percentage of nets are predictable than the interconnects or connections, more information about the distances between cells can be gained by prediction of connections or interconnects.

Microscopic Models Overview

The current state of the art in microscopic wire length prediction models can be classified into two different categories based on the output provided by the model. The first category provides length values for each individual wire as the output, based on a set of attributes extracted form the netlist. The models in the second category do not provide an exact value for the length of the wire, but instead they provide an ordering of the wires from the shortest to longest based on a metric. The individual wire length model developed by Bodapati and Najm [22] is a good example of the former while the mutual contraction model developed by Hu and Marek-Sadowska [23] is a good example of the later. The above two models are reviewed in the following section to understand the limitations to microscopic prediction inherently present in the modeling methodology.

Bodapati and Najm model

Bodapati and Najm developed a model to predict the routed length of each individual net. The underlying concept behind their method is to build a linear regression based model, characterized for a given place and route tool. The model takes as its input several metrics extracted from the structural attributes of the netlist and physical cell characteristics. It then predicts the length of the nets in this new netlist when placed using the characterized place and route tool. The model is built using the following attributes

classified as local and global parameters. The global parameters are parameters that are constant across the entire netlist and are listed below:

Number of cells in the design: N_c

Number of 2-pin nets in the design: N_{2agg}

Number of 3-pin nets in the design: N_{3aggt}

Number of 4-pin nets in the *design*: N_{4agg}

Number of 5-pin nets in the *design*: N_{5agg}

Number of 6 or more pin nets in the *design*: N_{6agg}

Average width of the cells in the design: W_{avg}

Height of the cell: H_c

Width of the core sites: W_{core}

Aspect ratio of the given design: R

Expected row utilization factor: U

The local parameters usually pertain to the net in question or its neighborhood, where the *neighborhood of a net n* is defined as the union of its *first level neighbors* and its *second level neighbors*. *First level neighbors* are the nets which are directly connected to the terminal cells of the net *n* and *Second level neighbors* are the nets which are connected to the terminal cells of the *first level neighbors*. The local parameters are:

Number of pins (terminals) on the net: P_{net}

Number of 2-pin nets in the *neighborhood of the net*: N_{2net}

Number of 3-pin nets in the *neighborhood of the net*: N_{3net}

Number of 4-pin nets in the *neighborhood of the net*: N_{4net}

Number of 5-pin nets in the *neighborhood of the net*: N_{5net}

Number of 6 or more pin nets in the *neighborhood of the net*: N_{6net}

Number of nets in the *neighborhood of the net*: N_{net}

The prediction model is developed as a function of a set of metrics derived from the global and local parameters, the details of which are explained in the following

paragraphs. The first metric is developed based on the number of pins in the net, since it is a very basic measure of how long the net is going to be. The number of pins of a net gives the number of cells connected to the net. Therefore, a base length L_{nbase} is calculated for every net as an average of the length of all the cells of a net placed horizontally or vertically adjacent to each other.

$$L_{nbase} = \frac{1}{2}\left(P_{net}H_{cell} + \frac{P_{net}W_{avg}}{U}\right)$$

(Eq4.4)
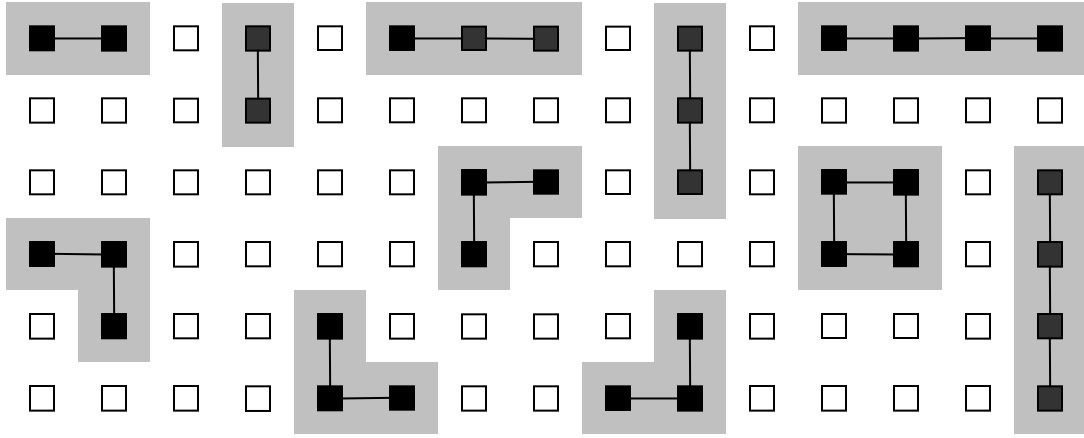
Figure 4.11 Possible placement configurations of 2, 3 and 4 pin nets

While Bodapati and Najm note that the wire lengths show a strong dependence on the base length of the nets, they also mention that nets with more than 7 pins have to be treated separately. Further they also observe that a majority of the 2 pin nets, 3 pin nets and 4 pin nets are placed such that their terminal cells are right next to each other as shown in Figure 4.11. Based on this, an estimate is made on all the possible positions of placing these lesser degree nets, and from this a degree of freedom measure is obtained for the cells connected to these nets. This degree of freedom measure is referred to as the congestion metric and is developed for 2 pin, 3 pin, 4 pin, 5 pin and 6 pin nets. The idea behind development of these metrics being, if the degree of freedom is large for say the 2 pin nets in the neighborhood of a given reference net, then the 2 pin nets will most likely be distributed over a large region. Since these 2 pin nets are observed to be short and

since the terminals of the given reference net are connected to these 2 pin nets, the terminals of the given net will also be distributed over the large region. As a result the reference net will be long.

In order to derive an expression for the congestion metrics, an expression for the number of rows in the given layout N_{rows} and, an expression for the number of core sites in the layout N_{core} are derived at first. The total length of the standard cell rows can be obtained using the two expressions as shown on either side of the equation below, from which the expression for N_{rows} is derived and subsequently an expression for N_{core} is derived as well.

$$\frac{N_c W_{avg}}{U} = N_{rows}(RN_{rows}H_c) \tag{Eq4.5}$$

$$N_{rows} = \sqrt{\frac{N_c W_{avg}}{H_c UR}} \tag{Eq4.6}$$

$$N_{core} = \frac{N_{rows}H_c R}{W_{core}} \tag{Eq4.7}$$

Assuming that the 2-pin nets can only be placed in one of the 2 configurations shown in Figure 4.11, the number of possible positions for a 2 pin net in the vertical configuration is P_{2cona} and in the horizontal configuration is P_{2conb}.

$$P_{2cona} = (N_{rows} - 1)U \frac{N_{core}}{W_{avg}} \tag{Eq4.8}$$

$$P_{2conb} = \left(\frac{N_{core}}{W_{avg}} - 1\right)UN_{rows} \tag{Eq4.9}$$

Now since the entire netlist is characterized by N_{2agg} 2 pin nets and since a given net neighborhood is characterized by N_{2net} 2-pin nets, the average number of positions available for the 2 pin nets in the given net's neighborhood can be approximately given by 2 pin congestion P_{2con}.

$$P_{2con} = \frac{(P_{2cona} + P_{2conb})N_{2net}}{N_{2agg}} \tag{Eq4.10}$$

Similarly expressions are derived for 3 pin congestion P_{3con}, 4 pin congestion P_{4con}, 5 pin congestion P_{5con}, 6 pin congestion P_{6con}. The expression for P_{3con} and P_{4con} is derived similar to P_{2con} except that the influence fewer pin nets are taken into account as well. For example for 3 pin congestion, this is done by calculating the proportion of nets of pin less than or equal to 3 in the neighborhood with respect to the total number of 3 or lesser pin nets. And the degree of freedom is estimated as directly proportional to this fraction, instead of just the fraction of just the 3 pin nets only. Also for P_{4con} only three possible placement configurations are considered.

$$P_{3con} = \frac{(P_{3cona} + P_{3conb} + P_{3conc})(N_{3net} + N_{2net})}{N_{2agg} + N_{3agg}} \qquad (Eq4.11)$$

$$P_{3cona} = \left(\frac{N_{core}}{W_{avg}} - 2\right) U N_{rows} \qquad (Eq4.12)$$

$$P_{3conb} = \left(\frac{N_{core}}{W_{avg}}\right) U (N_{rows} - 2) \qquad (Eq4.13)$$

$$P_{3conc} = 4\left(\frac{N_{core}}{W_{avg}} - 1\right) U (N_{rows} - 1) \qquad (Eq4.14)$$

$$P_{4con} = \frac{(P_{4cona} + P_{4conb} + P_{4conc})(N_{4net} + N_{3net} + N_{2net})}{N_{2agg} + N_{3agg} + N_{4agg}} \qquad (Eq4.15)$$

$$P_{4cona} = \left(\frac{N_{core}}{W_{avg}} - 3\right) U N_{rows} \qquad (Eq4.16)$$

$$P_{4conb} = \left(\frac{N_{core}}{W_{avg}}\right) U (N_{rows} - 3) \qquad (Eq4.17)$$

$$P_{4conc} = \left(\frac{N_{core}}{W_{avg}} - 1\right) U (N_{rows} - 1) \qquad (Eq4.18)$$

For P_{5con} and P_{6con} no unique placement configurations are considered. Instead an approximate estimate of the total number of positions available is obtained by deducting the number of possible sites taken away by 4 or lesser pin nets from an approximate estimate of the total number of possible positions for a single cell.

$$P_{5con} = \left(\frac{N_{core}}{W_{avg}} UN_{rows} - N_{2agg} - N_{3agg} - N_{4agg} \right) \frac{N_{5net}}{N_{5agg}} \qquad \text{(Eq4.19)}$$

$$P_{6con} = \left(\frac{N_{core}}{W_{avg}} UN_{rows} - N_{2agg} - N_{3agg} - N_{4agg} \right) \frac{N_{6net}}{N_{6agg}} \qquad \text{(Eq4.20)}$$

Finally, in order to account for the positions taken away from one 2 pin net by other 2 pin nets the metric N_{2oth} is derived as a fraction of the total number of 2 pin nets that are outside the neighborhood of the given net, where this fraction is given by the ratio of the number of nets in the neighborhood to the total number of nets in the design.

$$N_{2oth} = \left(N_{2agg} - N_{2net} \right) \frac{N_{net}}{N_c} \qquad \text{(Eq4.21)}$$

The length prediction model is then developed as a third order polynomial function of all of the above metrics. But then, the cross product terms that do not include L_{nbase} are ignored from this function to reduce the complexity. The authors report that this reduced the total number of terms in the equation by half to 20 without significantly affecting the quality of the fit. The coefficients of the polynomial are then derived by using the least squares fit of the equation to the actual lengths of nets extracted from a placed and routed design using a given place and route tool.

$$L_{net} = f \left(L_{nbase}, P_{2con}, P_{3con}, P_{4con}, P_{5con}, P_{6con}, N_{2oth} \right) \qquad \text{(Eq4.22)}$$

The above model however was observed to not work well for the nets of degree greater than 7. Therefore, a different model was developed for nets of degree greater than 7. It was based on estimating the dimensions of the bounding box of these nets and using it to approximate the value of the net length. First, the bounding box size is estimated by assuming that the number of cells inside the bounding box (N_{box}) will include all the cells in the neighborhood of the net, and in addition other 2 pin and 3 pin nets in the circuit proportional to the fraction of the total number of nets in the neighborhood with respect to the total number of cells in the design. The fraction of 2 pin nets to be included in N_{box}

is already given by N_{2oth}, and that of 3 pin nets given by N_{3oth} is estimated using a similar expression.

$$N_{3oth} = \left(N_{3agg} - N_{3net}\right)\frac{N_{net}}{N_c} \qquad \text{(Eq4.23)}$$

$$N_{box} = N_{net} + N_{2oth} + N_{3oth} \qquad \text{(Eq4.24)}$$

The dimension of the bounding box is then estimated based on the assumption that each cell in a net of degree greater than 7 will be placed in a different row, and therefore the height of the bounding box will at least span the minimum of P_{net} or N_{rows} rows. Given the number of cells inside the bounding box N_{box} the area of the bounding box can be estimated, and given the the height of the bounding box, the width of the bounding box can be estimated as well. Finally, the net length is estimated as the Rectilinear Steiner Minimal Tree length using the model from [32] based on the dimensions of the bounding box H_{bbox}, W_{bbox}.

$$if\left(\frac{W_{bbox}}{H_{bbox}} < 1\right) : L_{net} = \sqrt{P_{net}W_{bbox}H_{bbox}}$$
$$else : L_{net} = W_{bbox} + H_{bbox} \qquad \text{(Eq4.25)}$$

It was shown in [22] that the model cannot be used to predict the length of shorter wires. However for longer wires, average errors of the order of 20% to 30% were observed. However, it should be noted that the circuits tested were very small by current standards with no more than a thousand gates and a thousand nets.

Mutual contraction

The mutual contraction model, developed by Hu and Marek-Sadowska [23], is a metric to calculate the contraction force between every pair of cells connected through a net in the netlist. The value of this metric is then used as a guide to predict the order of connections from shortest to longest. But, the model is especially useful for predicting the shortest connections in the netlist because a pair of cells connected through a high mutual contraction connection tends to be short.

The model works by first estimating a weight for each connection $w'(c)$ as shown in Eq4.26, where n is the net from which the connection c is modeled and $d(n)$ gives the degree of the net.

$$w'(c) = \frac{2}{d(n) \cdot (d(n) - 1)}$$

(Eq4.26)
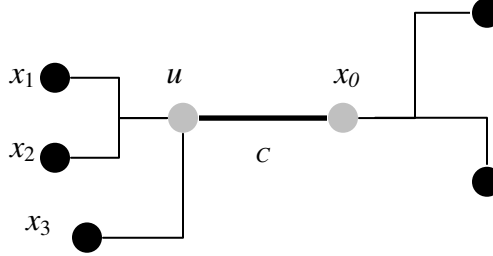
Figure 4.12 Illustration for relative weight estimation

The model then estimates the relative weight of a connection with respect to all the connections that are connected to one of its terminal nodes as shown in Eq4.27. In Eq4.27, u and x_0 represents the terminal cells of the connection c, and x_i represents the cells connected to u through a connection (see Figure 4.12).

$$w_r(u, x_0) = \frac{w'(u, x_0)}{\sum_{\forall x_i} w'(u, x_i)}$$

(Eq4.27)

This relative weight estimates a ratio of the weight of the reference connection c between nodes u and x_0 with respect to the sum total of the weights of all the connections connected to u. Finally the mutual contraction of a connection is estimated as the product of the relative weight of the connection with respect to each of its two terminal nodes.

$$mc(u, v) = w_r(u, v) \cdot w_r(v, u)$$

(Eq4.28)

Therefore, given the mutual contraction value for all the connections, the connections could be sorted in descending order based on their mutual contraction value. This sorted list is then used as a prediction of the connections length from the shortest to the longest. In fact, it was shown in [23] that the mutual contraction model performs

better in identifying the shorter connections than other models such as connectivity [51] and edge-separability [52].

Limitations of Microscopic Models

The individual length of a wire at the end of a placement is dependent on a number of factors that are based on the netlist and the placement process. The mutual contraction model, however, aims to make a prediction based on a simple metric that takes into account just the local neighborhood of the connection. This could be the source of a potential limitation of the mutual contraction model. However, this is not to say that the mutual contraction model does not add any value. On the contrary, the value and information gained though mutual contraction model can be combined with other models to gain additional insight into prediction. Further, the usefulness of the model has already been proven through its use in applications such as wire length driven placement [23] [53], timing driven placement [54] and technology mapping [55].

Figure 4.13 shows a plot of the cumulative length of the x number of shortest connections as predicted by the mutual contraction model and the connectivity model against the number of connections x for the IBM01 benchmark circuit. The lengths used in the plot are extracted from a simulated annealing based placement result. Because these models are more suited for predicting the shortest connections, the number of connections shown in the plot is limited to twice the number of cells in IBM01 benchmark circuit corresponding to the shortest connections in the netlist. It can be seen from this plot that, the mutual contraction model outperforms the connectivity model in [51].

Figure 4.13 also gives the cumulative length of the *actual* x shortest connections against the x number of shortest connections. It can be seen that this cumulative length of the actual shortest connections is much smaller than the cumulative length of the shortest

connections predicted by the mutual contraction model. This indicates that there is a vast scope for improvement in the model.
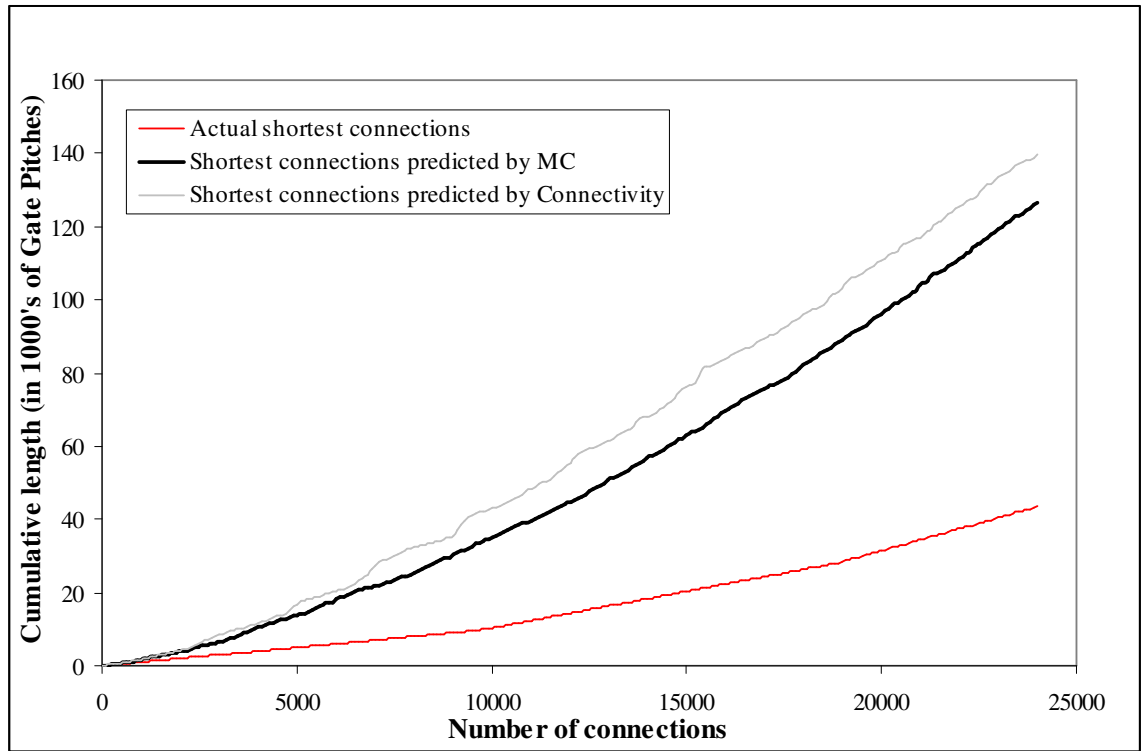
Figure 4.13 Cumulative length of shortest connections for IBM01

In fact, the percent difference in the cumulative length between the shortest connections predicted by mutual contraction model and the actual shortest connections is calculated for five different IBM benchmark circuits, and is shown in Table 4.1. For a netlist with G gates, this percent difference in cumulative length is calculated for the first G and 2G shortest connections. It can be seen from this table that the ideal prediction of the order of the connections from the shortest to the longest has a 78% smaller cumulative length than the mutual contraction model after the first G and 2G connections. Based on this quantitative evaluation, it is possible to conclude that the current model is in fact quite far of from optimal.

TABLE 4.1

DIFFERENCE IN CUMULATIVE LENGTH (MUTUAL CONTRACTION VS. ACTUAL SHORTEST)

Circuit	Percent Difference in cumulative length	
	After first G^* connections	After first $2G^*$ connections
IBM01	68.11	65.71
IBM02	77.51	80.01
IBM03	78.31	79.34
IBM04	80.34	78.42
IBM05	85.97	87.18
Average	78.05	78.13

* G represents the number of gates/cells in a given netlist

Unlike the mutual contraction model, the model by Bodapati and Najm takes into account both global parameters and local parameters. Thus more information is brought into the model to be used for prediction. In spite of the additional information, it was shown in [22] that the model still cannot be used for prediction of the length of short wires. Further, it assumes that there exists a polynomial relation between the length of the wires and factors derived from the global and local parameters. This could be a potential limitation of this model. For example, several relations might exist between the lengths and the various factors that control the lengths, and each of that relation might be valid only under a certain condition. In such a scenario prediction accuracy might be lost by using a single polynomial expression. In fact, the model already acknowledges the presence of multiple relationships, between length and prediction metrics, by using different prediction models for nets with degree less than 7 and nets with degree greater than or equal to 7.

Summary

An individual wire length could be predicted with reasonable accuracy only if the length of the wire does not vary widely from one optimal placement solution to another. Therefore an investigation is performed to evaluate this variability in individual wire lengths from one placement solution to another independent of placement tool. Based on these investigations it can be concluded that microscopic prediction is possible only for a

subset of the wires, corresponding to nearly 40% of nets or 20 to 35% of interconnects or 10 to 20% of connections in a given design. It is also shown that most of the predictable wires are from nets of smaller degrees and are short in nature.

Besides the limitations from variability, microscopic prediction accuracy can also be limited due to the modeling methodology. Most of the current models assume that individual wire length could be predicted using a single metric. For example, Bodapati and Najm's model uses a single complex polynomial function to predict the longer nets and takes as input several parameters derived from a given netlist, and mutual contraction model aims to predict the order of the connections from the shortest to longest solely based on a simple mathematical expression that takes as its input the degree of a connection's net and its neighboring connection's net degrees. However placement of gates is a complex process driven by the connectivity of the entire system. The resulting wire lengths may be a result of a number of different parameters that cannot be combined into a single mathematical expression. Consequently this strategy to model them using a single metric could be a bottleneck in prediction accuracy.

CHAPTER 5

NEW MICROSCOPIC MODEL

Accurate individual wire length prediction would be the best possible wire length prediction because it would provide an accurate physical perspective of the wires in the system. However it was shown in Chapter 4 that only a fraction of the wires can be predicted because the length of the remaining wires vary widely from one placement solution to another. But it was also shown in Chapter 4 that there exists a large scope for improvement in the individual wire length prediction model over current models, such as such as mutual contraction.
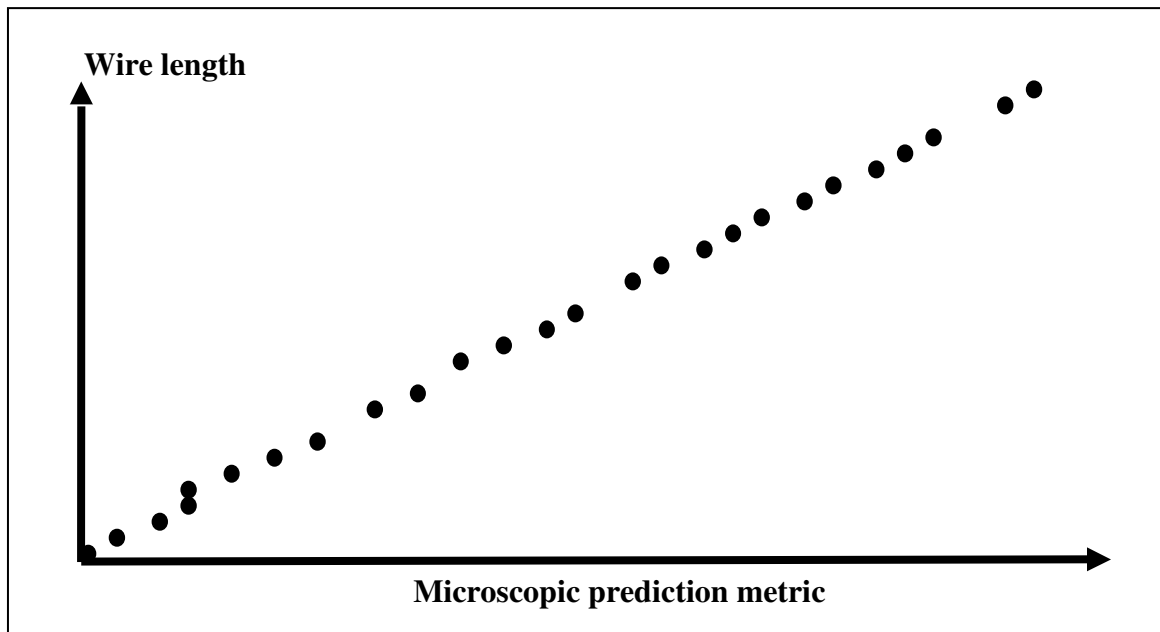
Figure 5.1 Ideal relation between microscopic prediction metric and wire length

For ideal prediction, the relation between the prediction metric and the actual lengths must appear as shown in Figure 5.1, where the length of the wire follows a perfect trend dependent on the metric. But in reality, what appears in such a plot between the actual prediction metrics and the wire length is as shown in Figure 5.2, where the colored dots represent the mispredictions.
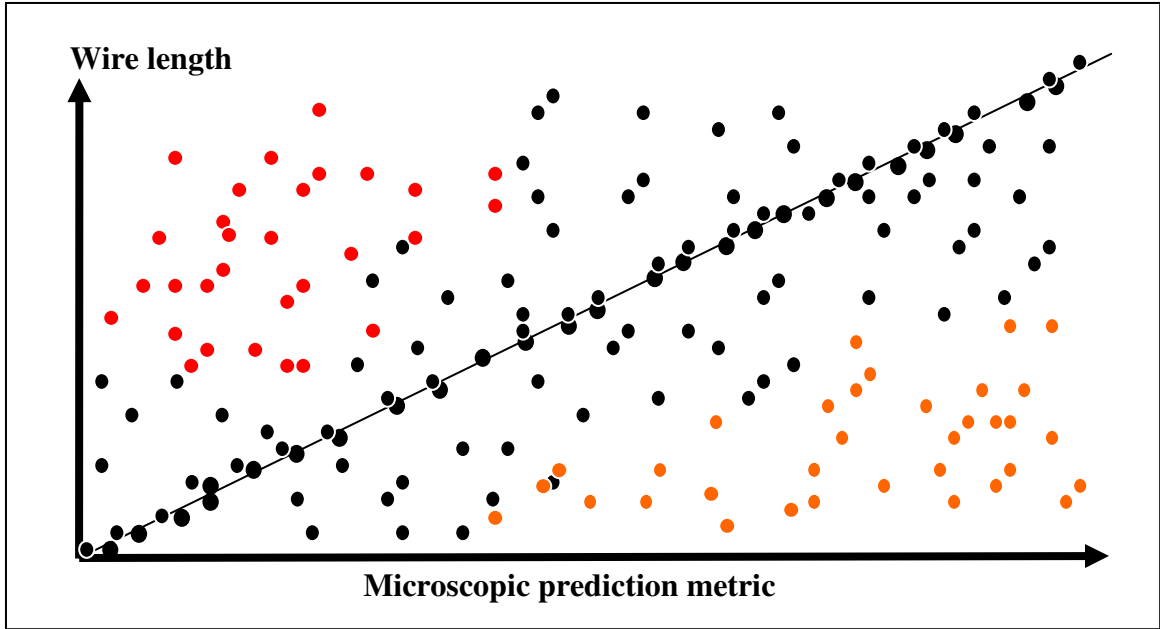
Figure 5.2 Relation observed between prediction metrics and actual wire length

For example, the connections in the IBM01 benchmark, ordered based on the mutual contraction metric, are divided among hundred bins with equal number of connections in each bin. The shortest connections predicted are in the first bin (bin 1) and the longest connections predicted are in the last bin (bin 100). Figure 5.3 shows the scatter plot of the length of the connections in each of these bins, which is ordered based on the mutual contraction metric. The figure also shows the average length of the connections in each of the bins. It can be seen from this plot that in fact the average length of the connections in each bin does increase from bin 1 to bin 100 with increasing mutual contraction. In spite of this, it is clear from the figure that there are a wide range of lengths in each of the bins. In fact the longest connection (length ~ 120 Gate Pitches) is predicted to be among the first 10% of connections, which ideally should include only the shortest 10% of connections. In spite of these errors, individual wire length prediction by mutual contraction model has already been used to effectively improve placement efficiency [22] [54], placement runtime [22], and synthesis [55]. Now if the huge mispredictions in the model could be reduced, or ideally eliminated, the effectiveness of

these applications could be improved and other advantages could be gained in the design process, by developing applications that take advantage of the improved accuracy.
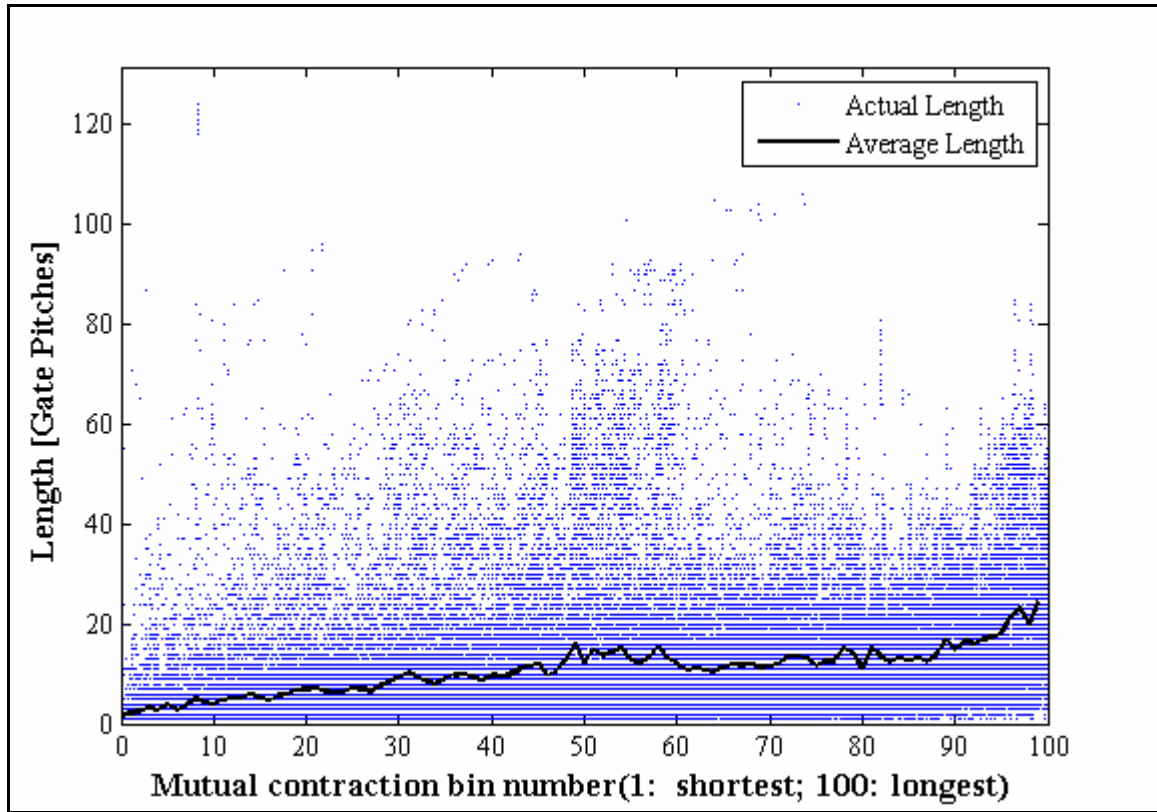
Figure 5.3 Connections sorted based on mutual contraction and arranged in bins

The objective in the development of a new microscopic prediction model is therefore to eliminate as many of these mispredictions as possible. It is shown in Chapter 4 that most of the wires that are predictable are very short in nature. Therefore, a new model is developed to predict the shorter wires more accurately than the mutual contraction model. The new model is based on several different prediction metrics, each of which can be used to identify a different group of shorter connections. They are combined together into a single model using a heuristic classification tree (HCT). Classification tree's are widely used in statistics for prediction purposes. The details of this model are presented in the first part of this chapter. The model is then used in a coarsening stage of a very simplified placement framework similar to the FPI framework in [53]. The details of this experiment are presented in the final part of the chapter.

New Heuristic Classification Tree

This new methodology of predicting shorter connections is based on identifying topological properties that result in shorter connections. The following methodology is adopted for the development and testing of the model. IBM01 benchmark is used as a sample circuit to study the structural properties of a netlist and their influence on the length of the connections. Those properties that are identified to make a connection short are then used to build the decision nodes of the new heuristic classification tree. The predictive ability of the model is then tested on the remaining seventeen benchmarks of the IBM benchmark suite.

The lengths of the connections for analyzing the properties are extracted from a simulated annealing based placement result. It is assumed that the length values extracted from simulated annealing based placement should be representative of the length achieved by the connections across different placement tools. The connection lengths are then used along with the topological attributes of the neighborhood of the connection to identify the properties that could cause the connection to be short.

Floating nodes

With current technologies the primary input or output (I/O) pins of the netlist need not be placed at the periphery of the chip. This is because the I/O terminals are not constrained to any fixed position and are free to move (float) about. For example, Figure 5.4 shows a sample graph where the vertices represent the cells in a netlist and the hyperedges represent the nets in the netlist. In this figure, vertex a is connected to only two other vertices, and one of them I/O1 is an input/output terminal that is not limited to a fixed position. Therefore vertex a will be pulled only in one direction, which is towards vertex b during placement. Because node a is free to move, it is a floating node. Since vertex b is connected to only two other vertices and since one of them (vertex a) is a floating node, vertex b will also be pulled only in a single direction i.e., towards vertex c.

Similarly vertex c will be pulled only towards vertex d, vertex d towards vertex e and so on. Consequently, nets (a,b), (b,c), (c,d), (d,e) and (e,f) will be short because there are no conflicting forces experienced by the nodes of these 2 pin nets. As a result, whenever a set of 2 pin nets are connected to each other in the form of a chain, and if one of them is terminated by an I/O pin as shown in Figure 5.4, the 2 pin nets can be assumed to be shorter in optimal placement solutions.
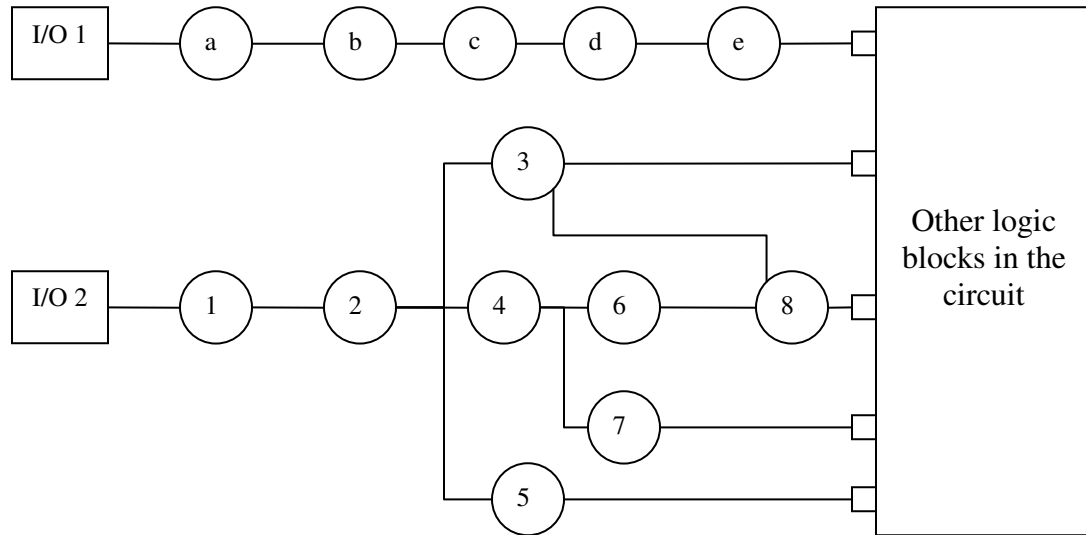
Figure 5.4 Direct and Indirect Paths

Based on this, the 2 pin net between vertices 1 and 2 in Figure 5.4 will also be short. However, the floating node concept cannot be extended to vertex 2. This is because vertex 2 is influenced by more than one another vertex, viz., vertices 3, 4 and 5, and each of these vertices may be pulled in a different direction by other vertices connected to them that are not free to move about. As a result, even if a vertex is connected to a floating node, if it is connected to more than 2 vertices including the floating node, it cannot be considered as a floating node. In other words, a floating node (FN) is defined as either an I/O pin or a vertex that is connected to only two nets, where both the nets are of degree 2 and one of them is connected to a floating node. This provides the first level of classification (*FN Prediction criterion*) in our new heuristic classification tree and is

shown in Figure 5.5. The details of the number of connections classified under each group and the average length of those connections identified in the group for IBM01 benchmark are also included in this figure. The data in the figure shows that 463 connections met the criteria and their average length was 1.15 Gate Pitches in IBM01.
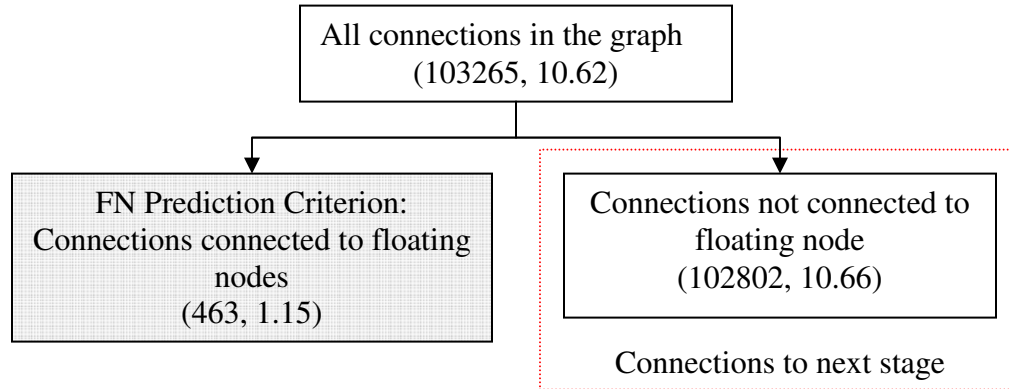
Figure 5.5. New Heuristic Classification Tree (Stage 1)

Alternative paths

An undirected graph can be generated from any netlist with vertices representing the nodes in the netlist and hyperedges representing the nets in the netlist. A closer look at such an undirected graph generated from IBM benchmark netlists reveals the presence of multiple paths between several pairs of nodes. Intuitively, in a wire length optimized placement, the nodes involved in multiple paths between them will have to be placed closer together. This is because by placing these nodes closer more than one wire's length is minimized. Further it was shown in Chapter 4 that most of the predicable connections which are short are from nets of smaller degrees. Based on this it is possible to hypothesize that whenever there are multiple paths between two nodes *and* whenever the paths involve lower degree nets it is highly probable that these two nodes will be placed close together in an optimized placement.

More generally, the paths between the two nodes could be either direct or indirect. A direct path is defined as a path that results from two nodes being part of the same net.

An indirect path is defined as a path that occurs when the two nodes are connected through a set of intermediate vertices and nets. For example, in Figure 5.4 there are two paths between vertices 3 and 4. The path through the connection (3, 4) in the hyperedge (2, 3, 4, 5) is a direct path. And the path from 3 to 4 through intermediate nodes 6 and 8 and intermediate hyperedges (4, 6, 7), (6, 8) and (8, 3) is the indirect path.

Direct paths

It is quite natural to assume that a direct path will result in more force between two nodes than an indirect path. This results in the second metric which relies upon the presence of multiple direct paths (MDP) between two nodes. The resulting classification tree is shown as the first split in Figure 5.6. It can be seen from this figure that, in the case of IBM01 benchmark placed using simulated annealing, the average length of connections between a pair of terminals connected with multiple direct paths is 7.36 gate pitches. This is in fact lesser than average length of connections between a pair of terminals connected with a single direct path.
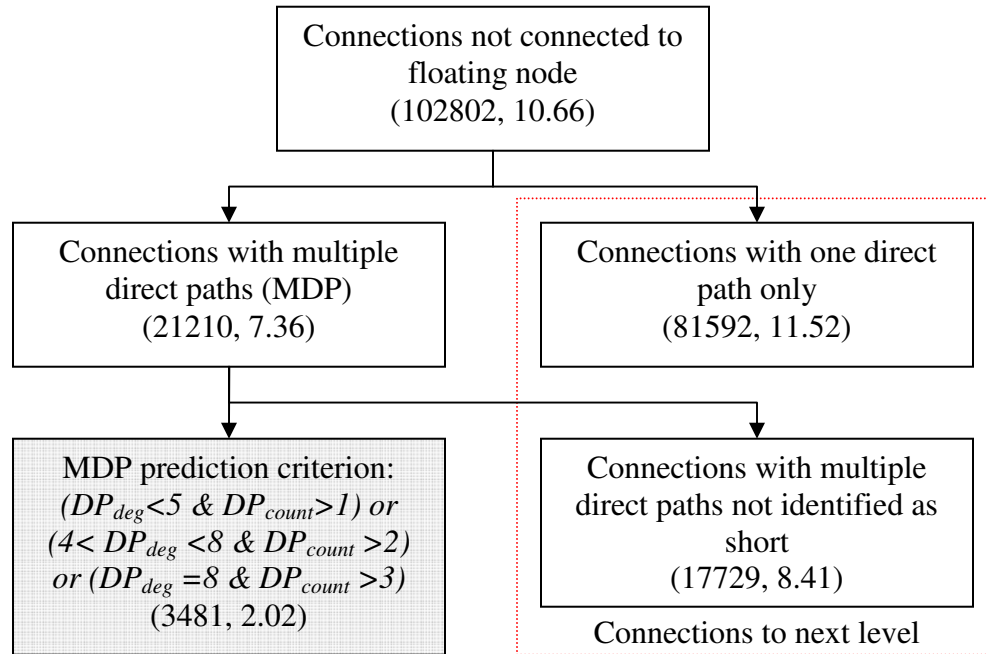
Figure 5.6. New Heuristic Classification Tree (Stage 2)
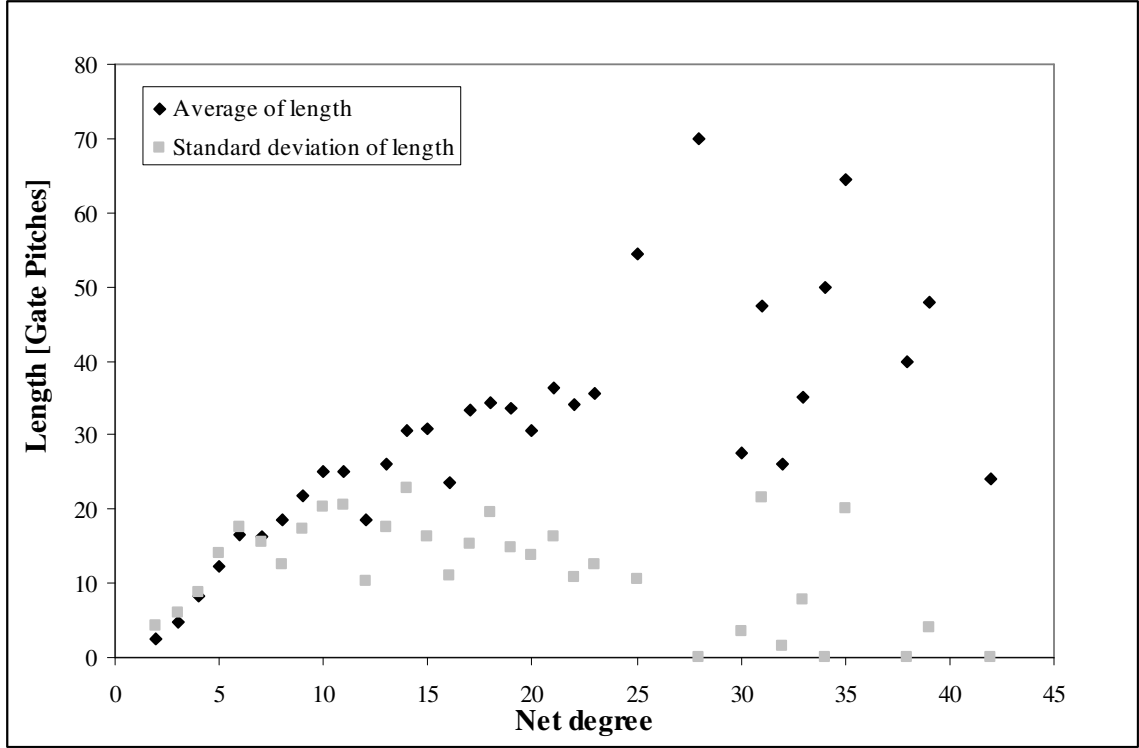
Figure 5.7 Average length of a net for different net degrees in IBM01

In a netlist, a net of lower degree generally has a smaller length than a net of higher degree. For example, Figure 5.7 gives the relation between the net degree and the average length of the nets of that degree in IBM01 benchmark circuit. Based on this it is possible to hypothesize that a lower degree net has a higher force of attraction between its cells. Consequently, when multiple direct paths exist in between a pair of terminals, the connections in the multiple direct paths are characterized by the lowest net degree of the connections that exist between the two terminals. The parameter used to represent the lowest net degree of the connections among the multiple direct paths is called the direct path degree (DP_{deg}).

Further, it is also possible to postulate that a pair of terminals will be placed closer together if there exists a large number of direct paths in between them. Therefore, the number of direct paths between the two terminals is also used to characterize the

105

connections in between the two terminals and the parameter direct path count (DP_{count}) is used to represent the number of connections between a pair of terminals.

Table 5.1 and Table 5.2 shows the average length and standard deviation for the connections characterized by these two parameters for IBM01. Data for connections of degree greater than 10 are excluded due to larger average lengths. It is clearly seen that, as expected, on average the connections with lower net degree and higher direct path counts are consistently short.

TABLE 5.1
AVERAGE LENGTH OF CONNECTIONS IN MULTIPLE DIRECT PATHS

DP_{deg}	DP_{count}					
	2	3	4	5	6	7
2	1.49	1.76	1.24	1	1	0
3	2.01	1.62	1.66	1.19	1	1
4	2.52	2.05	2.01	2	1	1.4
5	3.2	2.26	2.15	1.56	1	1
6	4.08	2.06	2.97	2.29	2	1
7	3.61	2.76	2.51	2.25	2.43	1.75
8	5.33	4.5	2.59	0	2	0
9	5.69	5.27	5.55	2	0	0
10	10.25	5.37	3.91	0	0	0

TABLE 5.2
STANDARD DEVIATION OF LENGTH OF CONNECTIONS IN MULTIPLE DIRECT PATHS

DP_{deg}	DP_{count}					
	2	3	4	5	6	7
2	1.37	2.17	0.49	0	0	0
3	1.48	1.09	1.64	0.47	0	0
4	3.19	1.45	1.19	1.22	0	0.49
5	4.42	1.6	1.32	0.5	0	0
6	4.1	1.26	1.94	0.88	0.82	0
7	3.17	1.85	1.59	1.09	0.9	0.83
8	4.19	3.36	1.47	0	0	0
9	5.42	5.89	5.78	1	0	0
10	10.53	4.2	2.6	0	0	0

Based on the Tables 5.1 and 5.2 certain groups of connections can be identified to be short depending on the values of the parameters DP_{deg} and DP_{count}. These groups are shown as shaded entries (both light and dark shades) in the tables. Connections that meet this criterion (*MDP prediction criterion*) is then used to predict the second set of short connections and is shown as the second decision split in the heuristic classification tree (stage 2) shown in Figure 5.6.

Further analysis on those connections identified as short in the IBM01 benchmark is carried out to evaluate the quality of the prediction. In this analysis, among the connections predicted as short, those with an actual length of 1 to 5 gate pitches are identified as correct prediction, those of length 6 to 10 gate pitches are identified as minor violation, those of length 11 to 50 gate pitches as moderate violation and those of length greater than 50 gate pitches as major violation. Given these definitions, the distribution of correct predictions, minor violations, moderate violations and major violations among those connections predicted as short is extracted as a function of the parameter DP_{deg}, and is given in Table 5.3. It is clearly seen that there are very few violations among those identified as short. In fact among the connections identified as short in the case IBM01 benchmark circuit, there are no connections of length greater than 50 gate pitches. Further, in total only 3.5% of the connections predicted as short are misclassified. This result can be viewed as the potential quality of the prediction possible in the second stage.

TABLE 5.3
DISTRIBUTION OF VIOLATIONS AMONG THE CONNECTIONS PREDICTED AS SHORT
(BASED ON A MULTIPLE DIRECT PATH CHARACTERIZING PARAMETER)

DP_{deg}	Correct Prediction	Minor Violation	Moderate Violation	Major Violation
2	507	6	2	0
3	1061	20	4	0
4	903	33	7	0
5	366	15	0	0
6	267	15	0	0
7	212	16	0	0
8	45	2	0	0

Indirect paths

While the previous section dealt with multiple direct paths between two nodes, this section deals with multiple paths, where one is a direct path and the others are indirect paths. Based on reasoning similar to that used in the case of multiple direct paths, it can be argued that whenever there are additional indirect paths (AIP) between two terminals, this could force the connections between the two terminals to be shorter under certain conditions.

There can be many indirect paths between two nodes and it will be time consuming to find all of them. Therefore, in this investigation, only the shortest indirect path with less than four nets along the path is considered for the creation of this heuristic classification tree. Because there can be many indirect paths between a pair of terminals connected by a direct path, and because each of the net in the indirect path can come from a different net degree and can be influenced by a number of other nets, it is necessary to have a figure of merit to characterize the connections in the path.

It was shown in Figure 5.7 that the lower degree nets tend to be shorter. Consequently, it can be assumed that a net of lower degree has a higher force pulling its cells together. Based on this assumption, a force weight is associated with each connection from a given net. A connection between the nodes (u, v) that is part of the net n of degree $d(n)$ is given a weight $f(u, v)$ given by the following expression.

$$f(u,v) = f(n) = \frac{2}{d(n)} \qquad \text{(Eq5.1)}$$

This metric assigns a larger weight to connections from smaller degree nets, with the weights never exceeding 1. Whenever there is more than one path present in between two nodes, it is possible to presume that the effect of the pulling force between the nodes will be influenced by all the connections involved in the path. A simple mathematical way to perform this function would be either to sum the weights of the connections in the path or to get a product of the weight of the connections.

A sum function would give a greater weight to a connection when the indirect path between the connections terminals has more nets in it. On the other hand, the product of weights function would give a lesser weight to a connection when the indirect path between the connections terminals has more nets in it. However, in an indirect path with a large number of nets, each of the number of nodes of the nets in the path will be subjected to a set of varied forces that determine their location. Consequently, it can be argued that the actual contraction force between the nodes in the path will be weaker when the indirect path has more nets. This will be contrary to the results produced by a sum function and concordant with the results produced by a product function.

Therefore, the product of weights of the nets that form the direct path (dp) and indirect path $ip1$ between the pair of terminals (u, v) is chosen to represent the combined weight of the paths $f_{total}(u, v)$ on the terminals (u, v). However if there is more than one indirect path, then the maximum of the combined weights of the paths is used as the metric as shown in Eq5.2.

$$f_{total}(u,v) = \max_{\forall j} \left(\prod_{n \in \{dp \cup ip_j\}} f(n) \right) \qquad \text{(Eq5.2)}$$

The number of unique weight values could be large since $f_{total}(u, v)$ is a fraction less than 1. Therefore, for the sake of easier analysis $f_{total}(u, v)$ is converted into an integer parameter called rank using the expression in Eq5.4.

$$Rank(u,v) = \left\lceil \frac{2}{f_{total}(u,v)} \right\rceil \qquad \text{(Eq5.4)}$$

The distribution of connections, average length of the connections and standard deviation of the length of the connections characterized by rank of the connection is shown in Table 5.4. It can be seen from this table that the average length of the connections and the standard deviation of the length of the connections increases with the rank of the connections. This agrees with the hypothesis that the more the nets in the

indirect paths, the smaller the combined weight $f_{total}(u, v)$, and hence longer the connections and greater the rank. Based on the values in Table 5.4, connections with a rank less than 14 are predicted as short and are shown shaded in the table. This criterion (*AIP prediction criterion*) is then used as the next prediction metric and the resulting heuristic classification tree is shown in Figure 5.8.

TABLE 5.4
STATISTICS OF LENGTH OF CONNECTIONS FOR EACH RANK

Rank	Number of connections	Average length	Standard deviation of length
3	246	1.63	0.95
4	666	1.64	1.75
5	1257	1.84	1.46
6	810	1.99	1.43
7	699	1.98	1.43
8	996	2.15	1.52
9	566	2.23	1.63
10	436	2.56	2.34
11	301	2.47	1.66
12	685	2.70	2.10
13	479	3.24	3.04
14	373	3.40	2.76
15	661	3.28	2.53
>16	24956	5.51	6.30

Given this set of connections predicted as short based on the AIP prediction criterion, they are further analyzed using the violation definitions that were used for analyzing the quality of prediction of the MDP prediction criterion. The distribution of violations among the connections predicted short using the AIP prediction criterion is extracted and is shown in Table 5.5 characterized by the rank of the connection. It is clear from this table that there are very few violations in the group of connections predicted as short based on presence of indirect paths. Again, similar to MDP prediction criterion, the AIP prediction criterion makes no major violations and results in less than 4% of misclassification in total. This result shows that, similar to the MDP criterion, the

new AIP prediction criterion could potentially make very accurate predictions of the shorter connections as well.
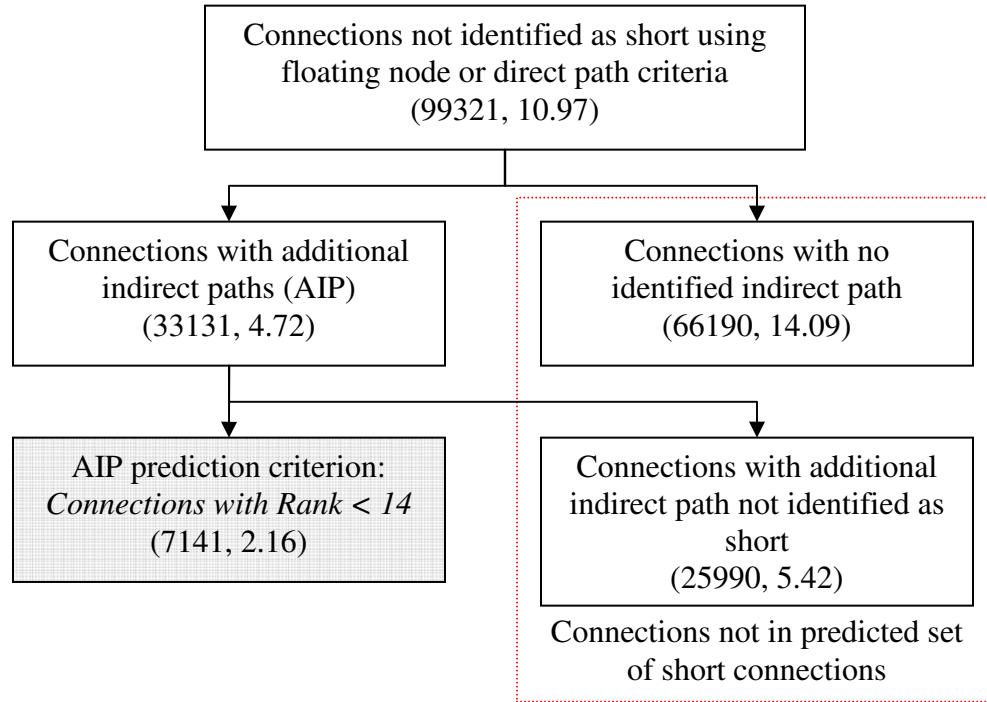
Figure 5.8. New Heuristic Classification Tree (Stage 3)

TABLE 5.5
DISTRIBUTION OF VIOLATIONS AMONG THE CONNECTIONS PREDICTED AS SHORT
(BASED ON ADDITIONAL INDIRECT PATH CHARACTERIZING PARAMETER)

Rank	Correct Prediction	Minor Violation	Moderate Violation	Major Violation
3	244	2	0	0
4	661	3	2	0
5	1217	34	6	0
6	790	18	2	0
7	687	10	2	0
8	961	31	4	0
9	545	16	5	0
10	406	25	5	0
11	284	17	0	0
12	640	38	7	0
13	423	47	9	0

Complete Heuristic Classification Tree Model

Figure 5.9 below shows the complete heuristic classification tree model for predicting the very short connections in a netlist.
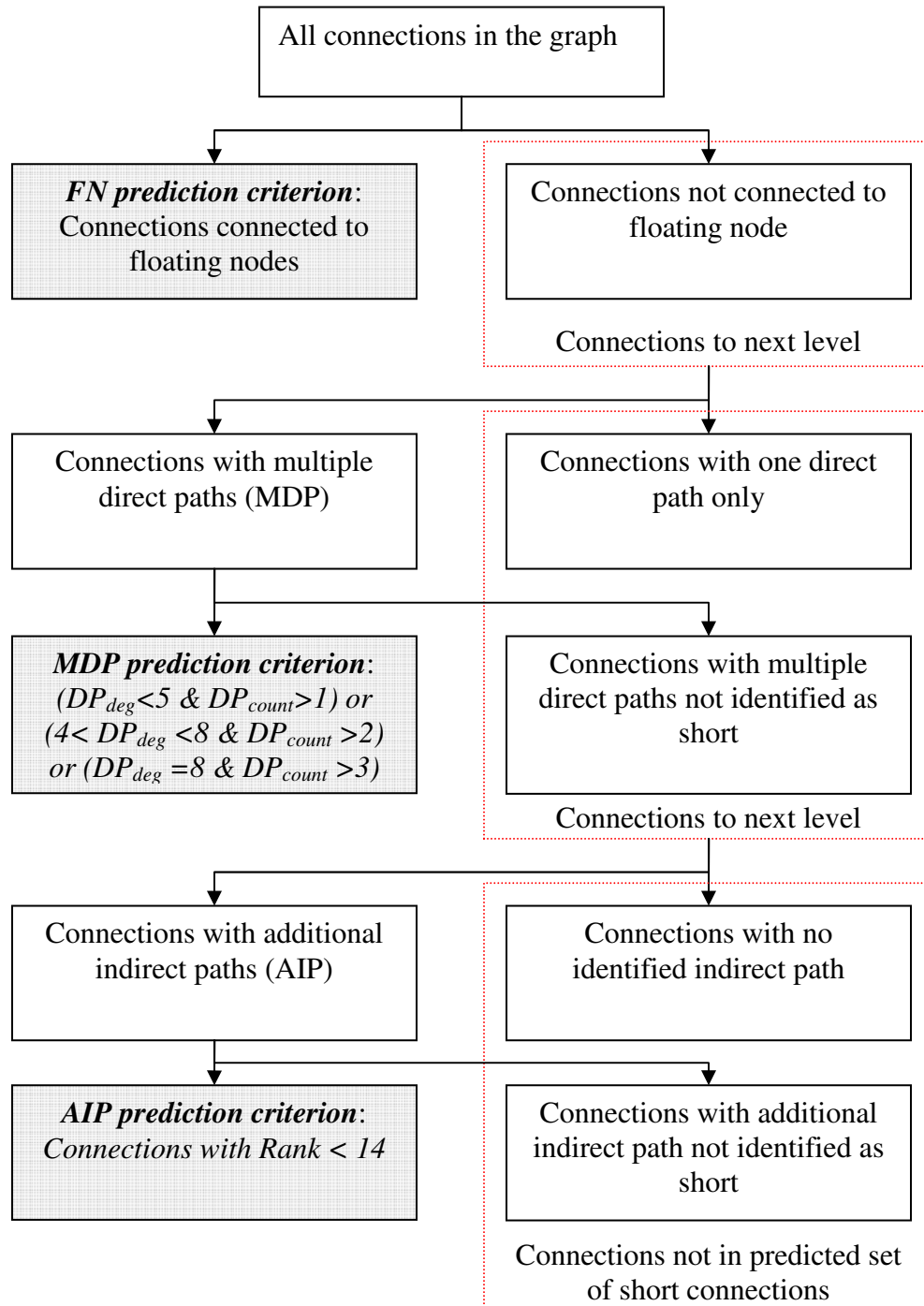
Figure 5.9 Complete Heuristic Classification Tree Model

New Model Results

The connections with the properties identified in the heuristic classification tree are arranged with the connections from the floating node at first followed by those with multiple direct paths and then those with an additional indirect path. The plot of cumulative length of these ordered connections produced by the new Heuristic Classification Tree is shown in Figure 5.10 and Figure 5.11 along with the cumulative length of the connections sorted by the mutual contraction metric. Although the model itself was tuned using the simulated annealing placement results of IBM01, the plot in Figure 5.10 is produced using the Dragon placement result of IBM16 benchmark and the plot in Figure 5.11 is produced using the Capo placement result of IBM18 benchmark. The plot shows the cumulative length of the connections predicted as short by the new model is smaller than the same number of shortest connections predicted by the mutual contraction model.
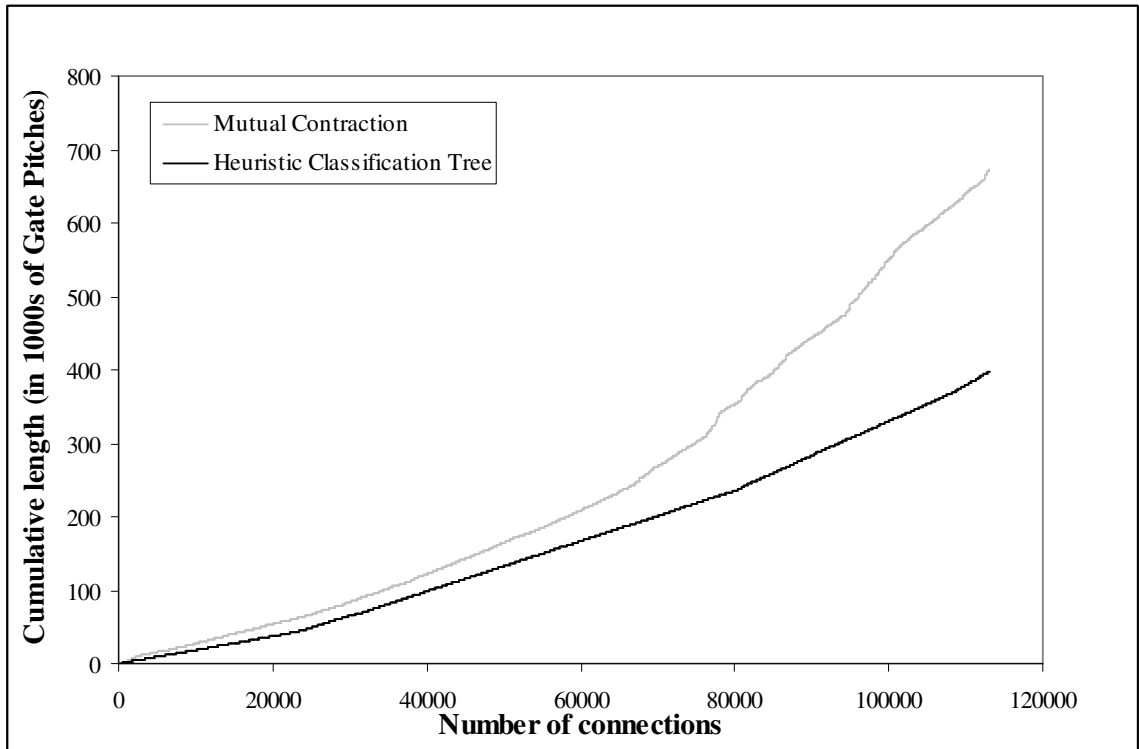
Figure 5.10 Cumulative length of connections from Dragon placement of IBM16
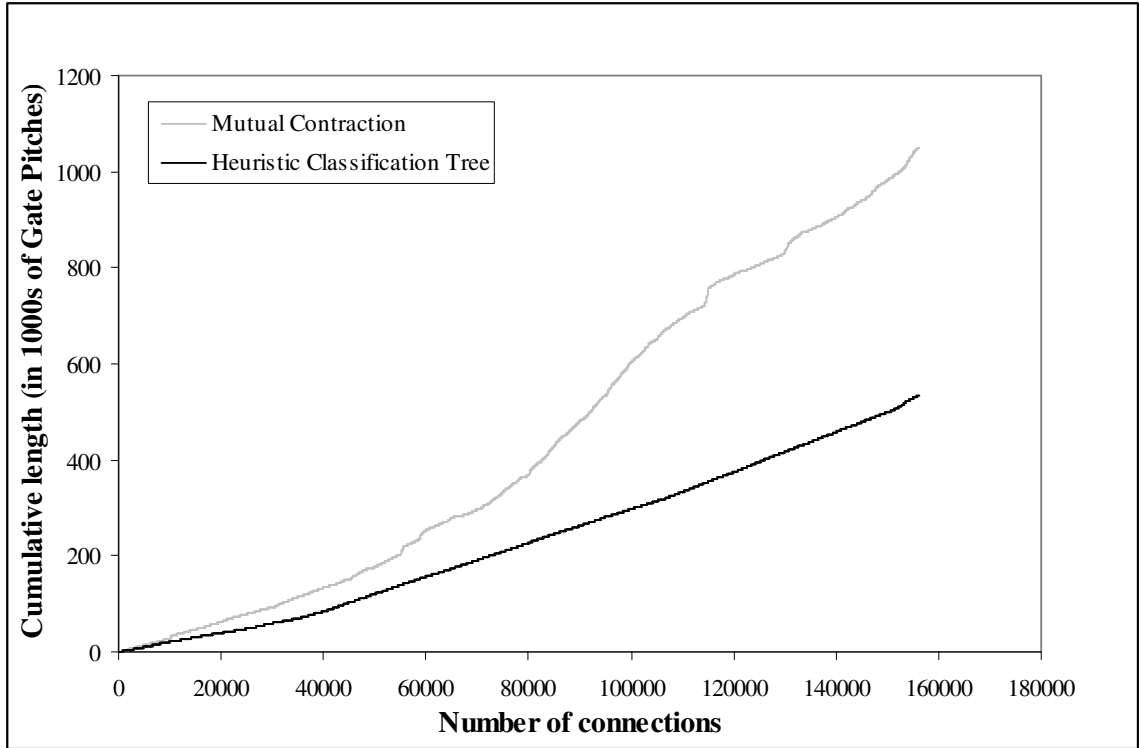
Figure 5.11 Cumulative length of connections from Capo placement of IBM18

A more quantitative comparison between the heuristic classification tree model and the mutual contraction model is provided in Table 5.6. The table provides the percent difference between the total length of the connections identified by the new heuristic classification tree with respect to the total length of the same number of short connections from mutual contraction model. The lengths of connections predicted by the new model are extracted from the placement results of three different placement tools viz., simulated annealing, Dragon and Capo, while the length of connections predicted by mutual contraction, used as the reference, is extracted only from the simulated annealing placement results. It can be seen from these results that the connections predicted as short by the new model have 28%, 27%, and 25% lesser total length than the connections identified by mutual contraction for the three different placement tools. However, it should be noted that results of IBM05 benchmark was not consistent with the other benchmarks. This could be attributed to the anomalous netlist properties of IBM05

114

benchmark which has a larger fraction of its connections derived from the higher degree nets, and has fewer connections that satisfy the heuristic classification prediction criteria. Excluding IBM05 results from the average for the percent cumulative length difference increases the average to 36%, 33% and 33% reduction for placement results from Simulated Annealing, Capo8.8 and Dragon respectively.

TABLE 5.6
HEURISTIC CLASSIFICATION TREE VS. MUTUAL CONTRACTION MODEL

Circuit	% Difference in cumulative length (with respect to mutual contraction)		
	SA	Dragon	Capo8.8
IBM02	-6.22	2.75	11.37
IBM03	-21.35	-6.45	-11.53
IBM04	-24.73	-15.81	-18.27
IBM05	87.17	100.76	71.58
IBM06	-25.23	-13.9	-15.74
IBM07	-40.22	-32.2	-35.59
IBM08	-38.12	-33.89	-33.66
IBM09	-10.88	-6	-7.99
IBM10	-49.74	-52.69	-49.34
IBM11	-35.03	-41.37	-37.13
IBM12	-43.23	-43.31	-35.64
IBM13	-30.25	-31.22	-32.18
IBM14	-40.01	-42.08	-45.8
IBM15	-48.33	-51.69	-54.99
IBM16	-46.85	-48.84	-51.23
IBM17	-55.7	-54.49	-54.98
IBM18	-55.79	-61.62	-60.97
Average	**-28.89**	**-25.49**	**-27.09**

An advantage of the heuristic classification tree is that it is easy to change the criteria used in the tree or add other new classification criteria to the tree to make the prediction more aggressive or to add other target prediction lengths. For example, an aggressive classification tree can be built by including only those groups identified by the darkly shaded regions in Tables 5.1 and 5.2. This classification has fewer mispredictions and fewer connections than the more conservative classification tree. In fact the cumulative length of the predicted shorter connections is less than those predicted by

115

mutual contraction by 50%, 44% and 38% with the placement results of simulated annealing, Capo and Dragon respectively.

The final analysis of the results is done to evaluate the quality of the prediction by estimating the number of mispredictions. For the sake of this analysis, a parameter called misprediction threshold $L_{threshold}$ is defined, such that if the actual length of a connection predicted as short is greater than this threshold, the connection is said to be mispredicted. Figure 5.12 shows the percentage of connections that are mispredicted for various values of the misprediction threshold in IBM12 benchmark using lengths from a simulated annealing placement. The plot shows the mispredicted data among the short connections predicted by the heuristic classification tree and the same number of shortest connections as predicted by the mutual contraction model. It can be seen from this figure that for any reasonable misprediction threshold the percentage of mispredictions in the new model is lesser than that observed in the mutual contraction model.
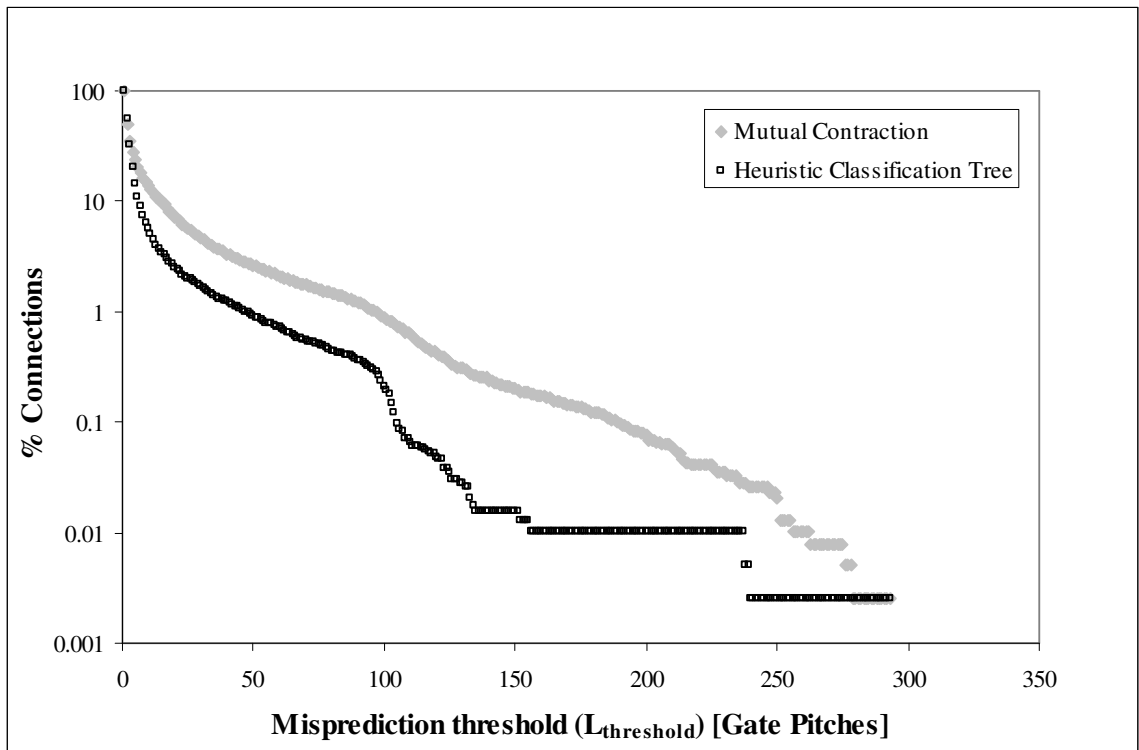
Figure 5.12 Difference in prediction quality observed in IBM16
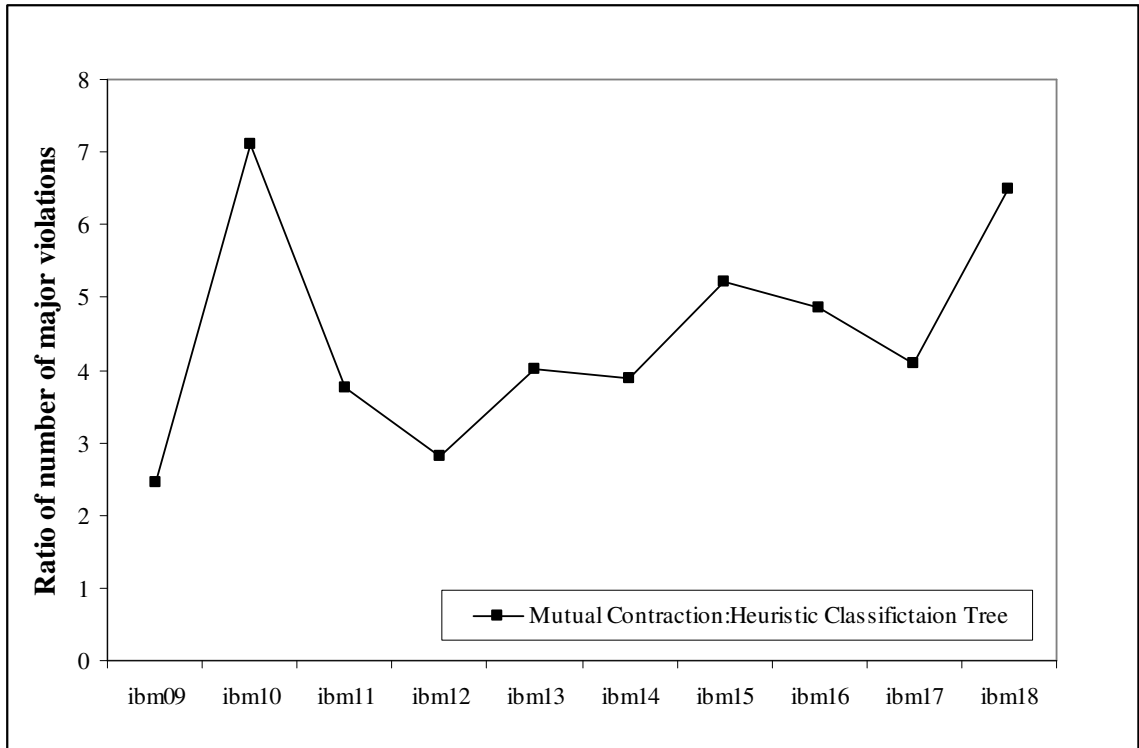
Figure 5.13 Ratio of major violation in the prediction

Results similar to that observed for IBM16 in Figure 5.12 were observed in other benchmarks as well and in shown condensed into a single plot in Figures 5.13 for the 10 largest benchmarks. Figure 5.13 shows the ratio of the number of major violations (mispredictions estimated with the misprediction threshold set to 50 gate pitches) among the predicted short connections in the mutual contraction with respect to the new heuristic classification tree model. It can be seen from this figure that the new model consistently outperforms the mutual contraction model and on an average has one fifth the number of major violations as a mutual contraction model. Based on these results it is possible to infer that the new heuristic model indeed improves the accuracy of the shorter connections prediction. However it should be mentioned that prediction using the new heuristic classification tree will take longer than the mutual contraction model. This is due to the fact that more information about the local neighborhood of a connection needs to be extracted to make a prediction with a heuristic classification tree.

New Model Application

One of the earliest applications of the mutual contraction model was its use in a multilevel placement framework FPI [53], where the mutual contraction model results are used to drive the coarsening decisions in a netlist. Given a netlist for placement, the idea was to first reduce the size of the netlist by coarsening and then apply a standard placement tool to the smaller coarsened netlist. Once the coarsened netlist is placed, it is un-coarsened and given to a detailed placer to further optimize the placement. Mutual contraction model was successfully used in this model along with a special coarsening algorithm to show effective reductions in the placement runtime.

```
┌─────────────────────────────────────────────┐
│  Predict short connections in the netlist using │
│         Heuristic Classification Tree         │
└─────────────────────────────────────────────┘
                        │
                        ▼
┌─────────────────────────────────────────────┐
│  Coarsen the netlist using the predicted short │
│                 connections                   │
└─────────────────────────────────────────────┘
                        │
                        ▼
┌─────────────────────────────────────────────┐
│    Place the coarsened netlist using Fengshui  │
│               placement tool                  │
└─────────────────────────────────────────────┘
                        │
                        ▼
┌─────────────────────────────────────────────┐
│  Un-coarsen the netlist to yield the original netlist │
│            with a global placement            │
└─────────────────────────────────────────────┘
                        │
                        ▼
┌─────────────────────────────────────────────┐
│  Optimize the global placement using Fengshui  │
│     detailed placer to yield the final placement │
└─────────────────────────────────────────────┘
```

Figure 5.14 Placement application framework of the heuristic classification tree

A placement application that utilizes the heuristic classification tree model in a similar manner is developed, and its framework is shown in Figure 5.14. In this framework, the heuristic classification tree model is used in the first stage to predict the shorter connections in netlist. The list of connections predicted as short is then used to reduce the size of the netlist. The coarsened netlist, which is smaller, is then fed to

Fengshui5.0 placement tool to find an optimal placement solution. After an optimal placement solution is found for the coarsened netlist, the cells in the coarsened netlist are un-coarsened to yield the original netlist. The cells of the original netlist are then assigned a position in the layout based on their parent cells in the coarsened netlist to yield a global placement for the original netlist. This global placement solution is then fed to the detailed placer of Fengshui5.0 placement tool to find the final optimal placement solution.

In this methodology, a very simple coarsening strategy is used to reduce the size of the netlist. The idea behind the coarsening step is that, since the connection predicted as short are most likely to be short in an optimized placement, the terminal cells of the connection will be placed closer together and therefore can be treated as a single cell. Therefore, if a connection is predicted as short, then the two terminal cells of the connection are merged into a single coarsened cell. However if one of the terminal cells is already merged into a different coarsened cell, then the number of cells in the coarsened cell is checked. If there are less than five cells in this coarsened cell, the cell that is not part of any coarsened cell is merged to the coarsened cell. In a case where both the cells are part of a coarsened cell, the two coarsened cells are merged if and only if a merger of the two coarsened cells will not create a coarsened cell that has more than 5 cells. In other words, the size of a coarsened cell is limited to 5 cells of the original netlist. The process is continued until all the connections that are predicted as short have been checked for possible coarsening. Once the cells are coarsened, the new netlist is created by replacing the original terminal cells of the nets with the coarsened cells. It is made sure that each net is connected to its coarsened cell through not more than one terminal. If all the terminals of the net are connected to the original cells within a single coarsened cell, the net is eliminated from the coarsened netlist. Thus the number of cells, number of nets and number of net terminals in the netlist are reduced by coarsening. Furthermore, the time taken to optimize a netlist is proportional to its size and therefore a

119

reduced netlist will take lesser time to optimize. After the coarsened netlist placement, the coarsened cells are replaced by the original cells that are part of it to yield the global placement solution of the original netlist.

The whole framework is implemented in C++, and its results in terms of wire length and run time is compared with the results produced by Fengshui5.0 with the original netlist as its direct input. These results are shown in Table 5.7 for 11 different IBM benchmarks. It can be seen from this table that the wire length increases marginally with the new framework. However, the run time was reduced by as much as 19%, and on an average the run time was reduced by nearly 10% accompanied by a 2% increase in wire length.

TABLE 5.7
PLACEMENT APPLICATION RESULT OF HEURISTIC CLASSIFICATION TREE

Circuit	Wire length in Gate Pitches			Runtime		
	New Framework	Fengshui	% Difference	New Framework	Fengshui	% Difference
IBM01	94977	93365	1.72	100.53	113.63	-11.53
IBM02	285573	270143	5.71	265.59	277.50	-4.29
IBM03	266170	265860	1.70	254.18	280.10	-9.25
IBM04	335008	328793	1.89	319.80	325.22	-1.66
IBM05	655584	628304	4.34	416.35	430.87	-3.37
IBM06	425186	425721	-0.12	432.56	474.83	-8.90
IBM07	619389	619647	-0.04	568.30	671.24	-15.33
IBM08	689410	673100	2.42	643.31	756.53	-14.96
IBM09	579720	573924	1.01	747.01	774.05	-3.49
IBM10	959854	943565	1.58	935.25	1042.85	-10.31
IBM18	3.466e6	3.339e6	1.97	2713.41	3339.14	-18.73

Further improvement in wire length should be possible by fine tuning the input control parameters given to Fengshui5.0 for detailed placement optimization and improving the coarsening strategy used. Although time taken by prediction and coarsening is a small fraction of the total placement, it should be possible to improve the runtime by optimizing the prediction, coarsening and uncoarsening stages. Additional improvements may also be possible if the framework is completely integrated into the placement tool.

Summary

Based on the results from previous investigations, a new microscopic prediction model is developed. The new model is developed to predict the shorter connections since earlier investigations revealed that the shorter wires are more predictable, and more information could be gained by predicting connections than interconnects or nets. Unlike the earlier models, the new model is designed as a set of several metrics organized in the form of a classification tree. Each node in the classification tree is a decision criterion and the connections that meet the criterion are predicted to be in a certain length range. The metrics are based on the connectivity of input/output terminals and the presence of additional paths between a pair of gates connected through a connection. The corner stones of this model are (i) Floating Node Criterion, (ii) Multiple Direct Paths Criterion and (iii) Additional Indirect Path Criterion. The resultant classification tree is much more accurate in predicting the shorter connections than the comparable mutual contraction model. In fact the new model has $1/5^{th}$ the number of major mispredictions than the mutual contraction model. Further the new model performs better than the mutual contraction independent of the placement tool used. In fact, the new model reduced the placement runtime by as much as 19% when used in a placement framework to perform coarsening.

CHAPTER 6

CONCLUSION AND FUTURE WORK

The primary objectives of the research performed as part of this thesis are (i) to carefully investigate the limitations to wire length prediction and (ii) identify opportunities and develop models to perform better wire length prediction. The results of the investigations performed to this effect were discussed in the previous chapters. Based on these investigations, this chapter presents the conclusion of this thesis and provides an overview of possible opportunities for future research work in the field of wire length prediction.

The chapter is organized into three main sections. The first section deals with macroscopic prediction, while the second deals with microscopic prediction. Each of the two sections provides a summary of the various conclusions that can be drawn form the investigations pertaining to the corresponding topic and future research work possibilities. Finally all of these results are tied together in the third and final conclusion section.

Macroscopic Prediction Summary

Macroscopic prediction refers to wire length prediction models that provide a global perspective on the wiring requirements of a given design. These models predict the wire length statistics, usually the wire length distribution. Most of the current wire length distribution models predict the less accurate interconnect length distribution because it is difficult to incorporate the more accurate net models into a wire length distribution prediction model. Although, ideally it is necessary to predict the post routing wire lengths, predicting them would be more difficult due to potentially larger variability as a result of the larger number of possible routed solutions. Therefore, the macroscopic

prediction research in this thesis is limited to a priori post-placement pre-routing interconnect length distribution prediction.

Interconnect length distribution is dependent on the actual final placement of the gates in the layout. Therefore, the length distribution will vary from one placement solution to another. This variation could be observed either between the placement results of the same placement tool (Intra-tool variability) or between the placement results from different placement tools (Inter-tool variability). This could consequently be a limiting factor to achievable prediction accuracy. Therefore, the variations from these causes were investigated. Based on these investigations it can be concluded that the intra-tool variability is generally less than the inter-tool variability. As a result, it is possible to hypothesize that the margin of prediction error could be reduced when a designer consistently uses the same placement tool. It was also shown that the variability for shorter wires is lesser than that of longer wires, and that the variability increases with a decrease in distribution. However, most often the longer wires are the sources of bottlenecks in a design. Therefore, when developing applications using predictions of longer wires, it is necessary to keep this variability into account. To aid in such scenarios, an empirical model that provides an idea of the variability of the distribution is also developed.

Accuracy of macroscopic wire length prediction could also be limited due to the modeling methodology used. Most of the current state-of-the-art wire length distribution prediction models calculate the distribution as a product of two functions viz., site density function and interconnect occupational probability function. The interconnect occupational probability values are most often calculated from a Terminal-to-Gate relation provided by the Rent's rule. The model developed by Davis, De and Meindl (DDM) is one of the very popular wire length distribution models. It calculates the interconnect occupational probability by applying the law of conservation of terminals to an imaginary set of three basic blocks A, B and C overlaid on the actual layout of the

placed netlist. Investigations were carried out to study the limitations to prediction accuracy due to the methodologies used in the model. It was shown that the DDM model overestimated the distribution of very short wires normally constituting 36% of the wires, by 94% on an average. Consequently, the distributions of the remaining wires were found to be underestimated. Further investigations revealed that the errors are due to the use of Rent's rule approximation and incorrect block sizes used in deriving the terminal counts.

Using the limitations to prediction accuracy identified from these investigations, a new interconnect length distribution model is developed as an improvement over the DDM model. The key corner stones of the new model are (i) a new model for parameters $p[ABC[l]]$ and $p[BC[l]]$, which are used in a modified Rent's rule to calculate the Interconnect-terminal-to-Block-size relation of a placed netlist, (ii) a look-up table that provides the p values of the two smallest block sizes used in model ($p[ABC[l=1]]$ and $p[BC[l=1]]$) and (iii) incorporation of the empirical variability model into the wire length distribution model. The result is a very highly accurate prediction of the distributions of the shorter wires. However the model has a poorer accuracy for the longer wires. This is due to the fact that the new model uses an approximation to model the parameter $p[BC[l]]$, which causes an over estimation of the difference in the terminal counts between blocks $BC[l]$ and $ABC[l]$ for very large length values. Consequently the length distribution of longer interconnects is overestimated.

Therefore, to improve the accuracy of macroscopic prediction in the current model, it is necessary to improve the model for $p[BC[l]]$. An alternative strategy would be to investigate possibilities for modeling the difference between the terminal counts of blocks $BC[l]$ and $ABC[l]$. However, ideally it is necessary to develop a priori post-routing wire length distribution model. Therefore, variability in the post-routing wire length distribution must be studied as well. Further, since the final actual wire length that influences the design specifications is based on a net model and not an interconnect model, a methodology to incorporate net models into wire length distribution models

must be investigated. This could result in much improved a priori post-routing wire length distribution models. Another area of possible future research is to develop a model by relaxing the assumptions used to build the current model, such as uniform square gate sizes and layouts.

Microscopic Prediction Summary

Microscopic predictions models provide a local perspective on the wiring requirements of the design. These models predict the length of each individual wires or predict the order of the wires from the shortest to the longest. Similar to macroscopic prediction, the microscopic prediction research performed as part of this thesis is limited to a priori post-placement pre-routing microscopic prediction.

An individual wire length could be predicted with reasonable accuracy only if the length of the wire does not vary widely from one optimal placement solution to another. Therefore an investigation is performed to evaluate this variability in individual wire lengths from one placement solution to another independent of placement tool. Based on these investigations it can be concluded that microscopic prediction is possible only for a subset of the wires, corresponding to nearly 40% of nets or 20 to 35% of interconnects or 10 to 20% of connections in a given design. It is also shown that most of the predictable wires are from nets of smaller degrees and are short in nature.

Although a larger percentage of nets are predictable, more information about the distances between cells could be gained by predicting connections or interconnections. Predictable interconnects could provide 40% more information than predictable nets, and predictable connections could provide 124% more information than predictable nets. These values are calculated by estimating the number of interconnects or connections of the predictable nets, and then comparing them to the number of predictable interconnects or predictable connections.

Besides the limitations from variability, microscopic prediction accuracy can also be limited due to the modeling methodology. Most of the current models assume that individual wire length could be predicted using a single metric. For example, Bodapati and Najm's model uses a single complex polynomial function to predict the longer nets and takes as input several parameters derived from a given netlist, and mutual contraction model aims to predict the order of the connections from the shortest to longest solely based on a simple mathematical expression that takes as its input the degree of a connection's net and its neighboring connection's net degrees. However placement of gates is a complex process driven by the connectivity of the entire system. The resulting wire lengths may be a result of a number of different parameters that cannot be combined into a single mathematical expression. Consequently this strategy to model them using a single metric could be a drawback to prediction accuracy.

Based on the results from these investigations, a new microscopic prediction model is developed. The new model is developed to predict the shorter connections since earlier investigations revealed that the shorter wires are more predictable, and more information could be gained by predicting connections than interconnects or nets. Unlike the earlier models, the new model is designed as a set of several metrics organized in the form of a classification tree. Each node in the classification tree is a decision criterion and the connections that meet the criterion are predicted to be in a certain length range. The metrics are based on the connectivity of input/output terminals and the presence of additional paths between a pair of gates connected through a connection. The corner stones of this model are (i) Floating Node Criterion, (ii) Multiple Direct Paths Criterion and (iii) Additional Indirect Path Criterion. The resultant classification tree is much more accurate in predicting the shorter connections than the comparable mutual contraction model. In fact the new model has $1/5^{th}$ the number of major mispredictions than the mutual contraction model. Further the new model performs better than the mutual contraction independent of the placement tool used. In fact, the new model reduced the

placement runtime by as much as 19% when used in a placement framework to perform coarsening.

Potential future research opportunities in the field of microscopic prediction include investigations of individual wire length variability with respect to timing instead of length. This could shed better light on the impact of the variability in timing. Currently the new microscopic model is limited to prediction of very short wires based on a set of three criteria. Future work could include research on additional criteria to the heuristic classification tree. This could increase the number of connections predicted, and could also help classify the connections into different length ranges. Further research could also be directed towards developing additional applications for the new heuristic classification tree model.

Key Knowledge Contributions

1. Inherent limitations to macroscopic prediction: For the first time, it is shown that the longer wires in a wire length distribution can have significant deviations for different placement runs and placement tools. The normalized standard deviation of the distribution of longer wires can be orders of magnitude larger than that of the short wires.

2. New macroscopic prediction model: A new wire length prediction model has been developed than can rapidly estimate the wire length distribution for a given netlist. In addition, this model includes the variations in the distribution. This model can be used to enhance system level simulators such as GENESYS and MINDS.

3. Inherent limitations to microscopic prediction: For the first time, the existence of a predictable set of wires that is independent of the placement run and placement tool is clearly identified. The predictable wires are usually short in nature and are from nets of smaller degrees.

4. New microscopic prediction model: A new HCT model has been developed that consistently predicts the shorter wires in a netlist. This new models shows 30-50% better prediction than current state-of-the-art models.

5. Application to microscopic prediction: It is shown that this new HCT model can be used to reduce placement time by up to 19% without significantly affecting placement quality.

Future Work

Wire length prediction is inherently difficult, limited by the modeling methodologies and the variability in wire lengths. Although models targeting macroscopic and microscopic prediction are developed separately in this thesis, more insight could be gained by combining these models into a single hybrid prediction model. For example, when the microscopic model predicts the length of shorter wires the length distribution of the shorter wires could be subtracted from the over all length distribution to provide the length distribution of the unpredicted wires. This could provide a better perspective of the wires whose individual lengths cannot be predicted directly. Thus hybrid wire length prediction models could largely improve the effectiveness of wire length prediction.

REFERENCES

[1] J. Cong, "An interconnect centric design flow for nanometer technologies," in *Proceedings of the IEEE*, vol. 89, pp. 505-528, April 2001.

[2] G. E. Moore, "Cramming more components onto integrated circuits," in *Proceedings of the IEEE*, vol. 86, pp. 82-85, January 1998.

[3] P. Saxena, N. Menezes, P. Cocchini, and D. A. Kirkpatrick, "Repeater scaling and its implication on CAD," in *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems,* vol. 23, pp. 451-463, April 2004.

[4] P. Kapur, G. Chandra, and K. C. Saraswat, "Power estimation in global interconnects and its reduction using a novel repeater optimization methodology," in *Proceedings of ACM Design Automation Conference,* 2002, pp. 461-466.

[5] W. Huang, M. R. Stan, K. Skadron, K. Sankaranarayanan, S. Gosh, and S. Velusamy, "Compact thermal modeling for temperature aware design," in *Proceedings of ACM Design Automation Conference,* 2004, pp. 878-883.

[6] W. E. Donath, "Placement and average interconnection lengths of computer logic," in *IEEE Transactions on Circuits and Systems,* vol. CAS-26, pp. 272-277, April 1979.

[7] B. S. Landman and R. L. Russo, "On a pin versus block relationship for partitions of logic graphs," in *IEEE Transactions on Computers,* vol. C-20, pp. 1469-1479, December 1971.

[8] P. Christie and D. Stroobandt, "The interpretation and application of Rent's Rule," in *IEEE Transactions on Very Large Scale Integration Systems*, vol. 8, pp. 639-648, December 2000.

[9] W. E. Donath, "Wire length distributions for placements of computer logic," in *IBM Journal of Research and Development,* vol. 25, pp. 152-155, May 1981.

[10] M. Pedram and B. Preas, "Accurate prediction of physical design characteristics for random logic," in *Proceedings of International Conference on Computer Design*, 1989, pp. 100-108.

[11] M. Pedram and B. Preas, "Interconnection length estimation for optimized standard cell layouts," in *Proceedings of International Conference on Computer-Aided Design*, 1989, pp. 390-393.

[12] T. Hamada, C. K. Cheng, and P. M. Chau, "A wire length estimation technique utilizing neighborhood density equations," in *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems,* vol. 15, pp. 912-922, August 1996.

[13] H. T. Heineken and W. Maly, "Standard cell interconnect length prediction from structural circuit attributes," in *Proceedings of Custom Integrated Circuits Conference*, 1996, pp. 167-170.

[14] J. A. Davis, V. K. De, and J. D. Meindl, "A stochastic wire length distribution for gigascale integration(GSI)-part I: derivation and validation," in *IEEE Transactions on Electronic Devices,* vol. 45, pp. 580-589, March 1998.

[15] D. Stroobandt, and J. V. Campenhout, "Accurate interconnect length estimations for predictions early in the design cycle," in *VLSI Design*, vol. 10, pp. 1-20, June 1999.

[16] D. Stroobandt, "A priori wire length distribution models with multiterminal nets," in *IEEE Transactions on Very Large Scale Integration (VLSI) Systems,* vol. 11, pp. 35-43, February 2003.

[17] P. Christie, "Rent exponent prediction methods," in *IEEE Transactions on Very Large Scale Integration (VLSI) Systems,* vol. 8, pp. 679-688, December 2000.

[18] P. Verplaetse, J. Dambre, D. Stroobandt, and J. V. Campenhout, "On partitioning vs. placement Rent properties," in *Proceedings of the ACM International Workshop on System Level Interconnect Prediction*, 2001, pp. 33–40.

[19] X. Yang, E. Bozorgzadeh and M. Sarrafzadeh, "Wire length estimation based on rent exponents of partitioning and placement," in *Proceedings of the ACM International Workshop on System Level Interconnect Prediction*, 2001, pp. 25-31.

[20] J. Dambre, D. Stroobandt, and J. V. Campenhout, "Towards the accurate prediction of placement wire length distributions in VLSI circuits," in *IEEE Transactions on Very Large Scale Integration (VLSI) Systems,* vol. 12, pp. 339-348, April 2004.

[21] A. Alvandpour, P. Larsson-Edefors and C. Svensson, "GLMC: Interconnect length estimation by growth-limited multifold clustering," in *Proceedings of IEEE International Symposium on Circuits and Systems*, 2000, pp. 465-468.

[22] S. Bodapati and F.N. Najm, "Pre-layout estimation of individual wire lengths," in *Proceedings of the ACM International Workshop on System Level Interconnect Prediction*, 2000, pp. 93-98.

[23] B. Hu and M. Marek-Sadowska, "Wire length prediction based clustering and its application in placement," in *Proceedings of IEEE/ACM Design Automation Conference*, 2003, pp. 800-805.

[24] A. E. Caldwell, A. B. Kahng, S. Mantik, I. L. Markov, and A. Zelikovsky, "On wirelength estimations for row based placement," in *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems,* vol. 18, pp. 1265-1278, September 1999.

[25] IBM Placement Benchmarks [Online] Available: http://er.cs.ucla.edu/benchmarks/ ibm-place/, January 2004

[26] Dragon3.01 Placement Tool [Online] Available: http://er.cs.ucla.edu/Dragon/ download.html, March 2005

[27] Capo8.8 Placement Tool [Online] Available: http://vlsicad.eecs.umich.edu/BK/ PDtools/tar.gz/Placement-bin/, March 2005

[28] Fengshui5.0 Placement Tool [Online] Available: http://vlsicad.cs.binghamton.edu/ software.html, April 2005

[29] mPL5.0 Placement Tool [Online] Available: http://cadlab.cs.ucla.edu/cpmo/, April 2005

[30] hMetis Partitioning Tool [Online] Available: http://www-users.cs.umn.edu/~karypis/ metis/hmetis/download.html, February 2004

[31] Rentc Rent Exponent Calculator [Online] Available: http://er.cs.ucla.edu/Dragon/ download.html, February 2006

[32] Semiconductor Industry Association, *International Technology Roadmap for Semiconductors*, 2001 [Online] Available: http://public.itrs.net/, June 2006

[33] A. Kahng, and S. Mantik, "Measurement of inherent noise in EDA tools," in *Proceedings of the IEEE International Symposium on Quality Electronic Design,* 2002, pp. 206-211.

[34] J.A. Davis, V.K De, and J.D. Meindl, "A stochastic wire length distribution for gigascale integration (GSI) Part II: applications to clock frequency, power dissipation, and chip size estimation," in *IEEE Transactions on Electron Devices*, vol. 45, pp. 590-597, March 1998.

[35] D. Stroobandt, and H. Van Marck, "Efficient representation of interconnect length distributions using generating polynomials," in *Proceedings of the ACM International Workshop on System Level Interconnect Prediction*, 2000, pp. 93-98.

[36] A. B. Kahng, S. Mantik, and D. Stroobandt, "Requirements for models of achievable routing," in *Proceedings of ACM International Symposium on Physical Design,* 2000, pp. 4-11.

[37] M. Y. Lanzerotti, G. Fiorenza, and R. A. Rand, "Assessment of on-chip wire-length distribution models," in *IEEE Transactions on Very Large Scale Integration (VLSI) Systems,* vol. 12, pp. 1108-1112, October 2004.

[38] A. Caldwell, Y. Cao, A. Kahng, F. Koushanfar, H. Lu, I. Markov, M. Oliver, D. Stroobandt, and D. Sylvester, "GTX: the MARCO GSRC technology extrapolation system," in *Proceedings of IEEE/ACM Design Automation Conference,* 2000, pp. 693–698.

[39] Y. Cao, C. Hu, X. Huang, A.B. Kahng, I.L. Markov, M. Oliver, D. Stroobandt and D. Sylvester, "Improved a priori interconnect predictions and technology extrapolation in the GTX System," in in *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 11, pp. 3-14, February 2003.

[40] D. Sylvester and K. Keuzer, "System-level performance modeling with BACPAC— Berkeley Advanced Chip Performance Calculator," in *Proceedings of the IEEE/ACM International Workshop on System Level Interconnect Prediction*, 1999, pp. 109–114.

[41] J.C. Eble, V.K. De, S.D. Wills and J.D. Meindl, "A generic system simulator (GENESYS) for ASIC technology and architecture beyond 2001," in *Proceedings of IEEE International ASIC Conference and Exhibit*, 1996, pp. 193-196.

[42] J. Dambre, D. Stroobandt, and J. Van Campenhout, "A probabilistic approach to clock cycle prediction," in *Proceedings of the ACM/SIGDA International Workshop on Timing Issues in the Specification and Synthesis of Digital Systems*, 2002, pp 9–15.

[43] J. Inoue, H. Ito, S. Gomi, T. Kyogoku, T. Uezono, K. Okada, and K. Masu, "Evaluation of on-chip transmission line interconnect using wire length distribution," in *Proceedings of the IEEE Asia South Pacific Design Automation Conference,* 2005, pp. 133-138

[44] P. Dasgupta, A. B. Kahng, and S. Muddu, "A novel metric for interconnect architecture performance," in *Proceedings of the IEEE Design, Automation and Test in Europe Conference and Exhibition,* 2003, pp.448-453

[45] W. Liao, and L. He, "Full-chip interconnect power estimation and simulation considering concurrent repeater and flip-flop insertion," in *Proceedings of the IEEE International Conference on Computer-Aided Design,* 2003, pp. 574-580

[46] X. Yang, R. Kastner, and M. Sarrafzadeh, "Congestion estimation during top-down placement," in *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems,* vol. 21, pp. 72-80, January 2002.

[47] K. Yamashita, and S. Odanaka, "Interconnect scaling scenario using a chip level interconnect model," in *IEEE Transactions on Electron Devices,* vol. 47, pp. 90-96, January 2000.

[48] Q. Chen, J. A. Davis, P. Zarkesh-Ha, and J. D. Meindl, "A compact physical via blockage model," in *IEEE Transactions on Very Large Scale Integration (VLSI) Systems,* vol. 8, pp. 689-692, December 2000.

[49] P. Christie, "A differential equation for placement analysis," in *IEEE Transactions on Very Large Scale Integration (VLSI) Systems,* vol. 9, pp. 913-921, December 2001

[50] J. Dambre, P. Verplaetse, D. Stroobandt and J. V. Campenhout, "On rent's rule for rectangular regions," in *Proceedings of the ACM International Workshop on System Level Interconnect Prediction*, 2001, pp. 49-56.

[51] S. Hauck and G. Borriello, "An evaluation of bipartitioning techniques," in *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 16, pp 849-866, August 1997.

[52] J. Cong and S. K. Lim, "Edge separability based circuit clustering with application to circuit partitioning," in *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 23, pp 346-357, March 2004.

[53] FPI2.0 Placement Tool [Online] Available: http://cornet.ece.ucsb.edu/~hu/fpi/ fpi2.0/, April 2005

[54] Q. Liu and M. Marek-Sadowska, "*Individual wire-length prediction with application to timing-driven placement*" in *IEEE Transactions on Very Large Scale Integration (VLSI) Systems,* vol. 12, pp. 1004-1014, October 2004.

[55] Q. Liu and M. Marek-Sadowska, "Wire length prediction based technology mapping and fanout optimization," in *Proceedings of International Symposium on Physical Design*, 2005, pp. 145-151.