Generalized Force Directed Relaxation with Optimal Regions and Its Applications to Circuit Placement *

(Invited Paper: A Tribute to Professor Satoshi Goto)

Yao-Wen Chang
Graduate Institute of Electronics Engineering, National Taiwan University, Taipei 106, Taiwan
Department of Electrical Engineering, National Taiwan University, Taipei 106, Taiwan
ywchang@ntu.edu.tw

ABSTRACT

This paper introduces popular algorithmic paradigms for circuit placement, presents Goto's classical placement framework based on the generalized force directed relaxation (GFDR) method with an optimal region (OR) formulation and its impacts on modern circuit placement and applications, and provides future placement research directions based on the GFDR and OR formulations.

Keywords

Physical Design, Placement, Iterative placement, Constructive placement, Nondeterministic placement, Mixed-size placement, Mixed-cell-height placement, Simulated annealing, Force-directed relaxation, Optimal region, FPGA, Routability, Timing

1. INTRODUCTION

The placement problem is to assign circuit modules (e.g., standard cells, macros, etc.) to desired positions on the chip, such that no two modules overlap with each other and some predefined cost metric (e.g., wirelength, routability, timing) is optimized; Figure 1 illustrates a placement instance, where interconnections among circuit modules are not shown. Placement is a major step in physical design, which plays a pivotal role in determining the final quality of a circuit design. As such, placement has been studied for several decades since the early days of integrated circuit designs. Recently, it has attracted much more attention than ever, due mainly to the dramatic growth in design complexity and many emerging technology challenges.

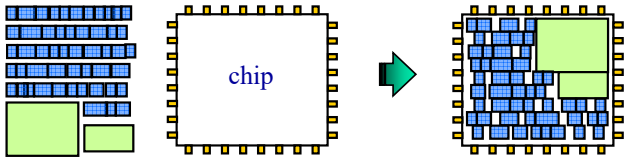

Figure 1: The circuit placement problem.

There are three major algorithmic paradigms for circuit placement [12, 15]: (1) constructive approach: once the position of a module is fixed, it is not changed anymore (for example, cluster growth, min-cut partitioning, quadratic placement, etc.), (2) iterative approach: intermediate placements are modified to improve the solution quality until some termination conditions are met (for example, force-directed method, etc.), and (3) nondeterministic approach: a placer may behave differently from run to run (say, based on a probabilistic formulation), even for the same input (for example, simulated annealing, genetic algorithm, etc.). The three types of approaches can be combined to further improve placement solutions; for example, an initial placement can be obtained by a constructive approach, followed by iterative improvement to enhance the placement, and further refined by simulated annealing to get the final placement.

Professor Satoshi Goto's 1981 paper [6] is an influential milestone work on circuit placement, which elegantly combines two algorithmic paradigms with a constructive initial placement followed by iterative improvement to obtain an effective and efficient placement framework. The initial placement, called *sub-optimum random generation (SORG)*, sequentially selects unplaced modules according to their connectivity to other modules and places them in desired positions to minimize the total wirelength. The iterative improvement, called *generalized force-directed relaxation (GFDR)*, repeatedly interchanges a set of modules to minimize the total wirelength.

Besides demonstrating an important breakthrough in placement techniques, it exemplifies a well-written, influential paper that combines both theoretical and empirical ingre-

*This work was partially supported by AnaGlobe, MediaTek, RealTek, TSMC, MOST of Taiwan under Grant NSC 102-2221-E-002-235-MY3, NSC 102-2923-E-002-006-MY3, MoST 103-2221-E-002-259-MY3, MoST 103-2812-8-002-003, MoST 104-2221-E-002-132-MY3, and NTU under Grant NTU-ERP-105R8951.

ISPD'17, March 19-22, 2017, Portland, OR, USA

DOI: http://dx.doi.org/10.1145/3036669.3038250

dients. Circuit placement is an NP-complete problem that had already been explored by many researchers in both industry and academia by 1981 [12, 15]. Professor Goto shows in his paper the then surprising theoretical finding that the *optimal region (OR)* for a module to achieve the minimum wirelength can be computed by an efficient median formulation. This is an exciting work where a solid theory and its practical benefits can be achieved simultaneously. As a result, its impacts is profound and long-lasting—the technique is not only well adopted in industry (like NEC), but is incorporated into even well-known modern placers such as FastPlace [11], mPL [1], NTUplace [3, 7], and POLAR [9].

In this article, we review Goto's OR- and GFDR-based classical placement framework and explain its impacts on modern circuit placement and applications. With their great success in modern placement, we further discuss potential future research directions based on the OR and GFDR formulations.

The rest of this paper is organized as follows: Section 2 reviews the OR formulation and the GFDR method, and also discusses their applications to modern placement. Section 3 explores future research directions based on the techniques presented in [6]. Finally, we conclude our work in Section 4.

2. OPTIMAL REGIONS AND GENERALIZED FORCE-DIRECTED RELAXATION

We first review the two key techniques presented in the pioneering work [6] and then address their applications to modern placement.

2.1 Optimal Region

For a set of nets and other modules connecting to a module m_i, the *optimal region* for m_i is defined as the region for placing m_i with the minimum total wirelength, with all the other modules being fixed. The optimal region can be found based on the median formulation proposed in [6]; see Figure 2 for an illustration. We follow the notations in [13] for easier presentation. Let $E_i = \{e_{i,1}, e_{i,2}, ..., e_{i,n_i}\}$ be the set of n_i nets connecting to m_i. Assume that pins are located in the centers of modules. For each net $e_{i,j} \in E_i$, we define the *bounding box* of $e_{i,j}$ as the minimum enclosing rectangle of all pins for $e_{i,j}$, excluding the module m_i. Let $x_{e_{i,j},l}$, $x_{e_{i,j},u}$, $y_{e_{i,j},l}$, and $y_{e_{i,j},u}$ be the respective left, right, lower, and upper boundaries of the bounding box of $e_{i,j}$. Let $\tilde{X}_i = \langle \tilde{x}_{i,1}, \tilde{x}_{i,2}, ..., \tilde{x}_{i,2n_i} \rangle$ be the sorted sequence of the x-boundaries $\{x_{e_{i,1},l}, x_{e_{i,1},u}, x_{e_{i,2},l}, x_{e_{i,2},u}, ..., x_{i,n_i,l}, x_{e_{i,n_i},u}\}$, and $\tilde{Y}_i = \langle \tilde{y}_{i,1}, \tilde{y}_{i,2}, ..., \tilde{y}_{i,2n_i} \rangle$ the sorted sequence of the y-boundaries $\{y_{e_{i,1},l}, y_{e_{i,1},u}, y_{e_{i,2},l}, y_{e_{i,2},u}, ..., y_{e_{i,n_i},l}, y_{e_{i,n_i},u}\}$. The optimal x and y coordinates of m_i can be obtained by solving the following optimization problem:

$$\min \sum_{j=1}^{2n_i} |x_i - \tilde{x}_{i,j}| + \sum_{j=1}^{2n_i} |y_i - \tilde{y}_{i,j}| . \quad (1)$$

The optimal solution for Equation (1) can be solved by finding the medians of $\tilde{X}_i$ and $\tilde{Y}_i$. Because both $\tilde{X}_i$ and $\tilde{Y}_i$ contain even members, both the medians of $\tilde{X}_i$ and $\tilde{Y}_i$ can be the optimal solutions, which form the left, right, lower, and upper boundaries for the optimal region of m_i.

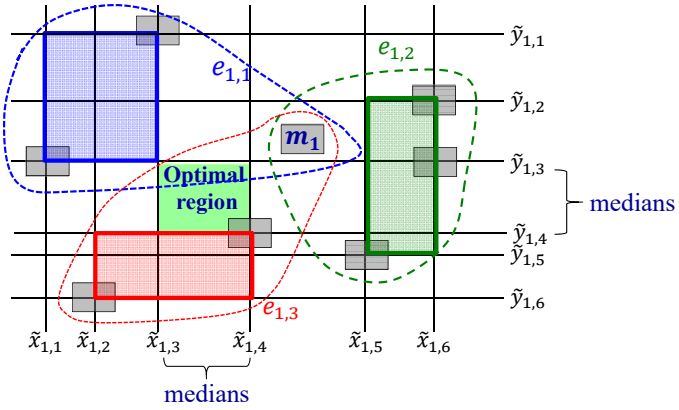

Figure 2: Optimal region computation for a module m_i, assuming that pins are located in the centers of modules.

2.2 Generalized Force Directed Relaxation

After an initial placement is obtained, the work [6] employs an iterative scheme to improve the objective based on the concepts of ϵ-neighborhood and λ-exchange. Fixing all other modules in their current positions, we can compute the optimal region for module m based on the formulation presented in Section 2.1. Suppose the optimal location of the module m is bin (s, t). Then, modules located at bin (i, j), where $|i - s| + |j - t| \leq \epsilon$, are called ϵ-neighbors of module m. For the example shown in Figure 3, the 1-neighbors of module A are $\{B, C, D, E, F\}$ if the optimal location of module A is occupied by module B.

Starting from a module m, GFDR computes the ϵ-neighbors of m, and for each ϵ-neighbor of m, GFDR further computes its ϵ-neighbors, and so on. For example, Figure 3 illustrates the exchange sequence with $\epsilon = 1$ and $\lambda = 3$ (three modules for such exchanges): $A \rightarrow B \rightarrow J \rightarrow A$ by moving module A to B's bin, module B to J's bin, and module J to A's bin. All module exchange sequences with $\epsilon = 1$ and $\lambda = 3$ are explored during the iterative improvement process, and the sequence with the minimum total wirelength is selected, or no exchange is performed if we cannot find an exchange sequence with a smaller wirelength.

2.3 Applications to Modern Placement

A modern chip could contain tens of millions of modules (standard cells). To handle this high design complexity, modern placement typically consists of three major stages: (1) global placement, (2) legalization, and (3) detailed placement. See Figure 4 for an illustration. Global placement computes the best position for each module to minimize the cost (e.g., wirelength), while ignoring module overlaps. Then, legalization places modules into desired positions and

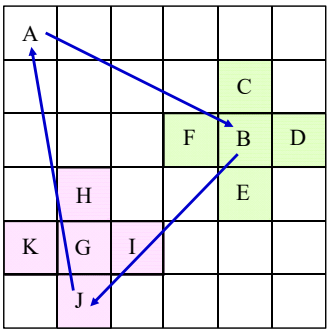

Figure 3: Trial interchange of modules with ϵ-neighborhood.

removes all overlaps among the modules. Finally, detailed placement further refines the module positions to obtain the final solution.

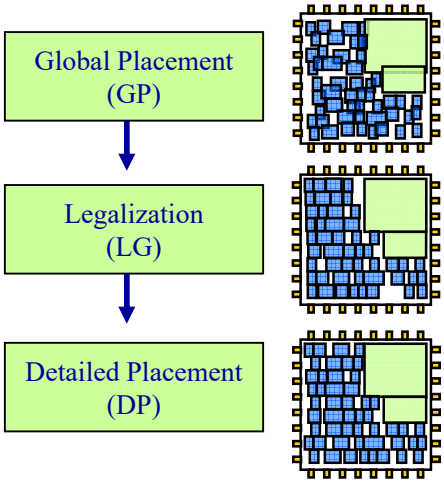

Figure 4: The modern placement flow typically consists of three stages: global placement, legalization, and detailed placement.

Even though the methods of ORs and GFDR were developed more than three decades ago, these methods are still pervasively used in modern placement algorithms. The following subsections just give several of such examples.

2.3.1 *Applications of the Optimal Regions*

FastPlace performs three types of module swaps to refine its placement solutions during its detailed placement stage: global swap, vertical swap, and local reordering [11]. The underlying idea of global swap is to find the optimal region for a module m_i in the placement region and swap m_i with another module m_j or a space in the optimal region to minimize the wirelength. A gain function is defined to choose the most profitable one among potential multiple modules or spaces in the optimal region.

Tseng, Chang, and Liu presented a detailed placement algorithm to handle the minimum-implant-area (MIA) constraint [13], where MIA-violating modules of the same threshold voltage are clustered in their optimal regions, and then an existing detailed placement algorithm is applied to deal with the cluster-based placement problem. This way the algorithm can simultaneously solve MIA violations while minimizing the wirelength. They extend the concept of the optimal region presented in [6] to placement with clusters of modules, called the *cluster-based optimal region*, to find the minimum wirelength region for a cluster. For a cluster $u_\ell = \{c_1, c_2, ..., c_{n_{u_\ell}}\}$ with n_{u_ℓ} modules, the optimization problem of finding the cluster-based optimal region is given as follows:

$$\min \sum_{i=1}^{n_{u_\ell}} \sum_{j=1}^{2n_i} |x_i - \tilde{x}_{i,j}| + \sum_{i=1}^{n_{u_\ell}} \sum_{j=1}^{2n_i} |y_i - \tilde{y}_{i,j}| \,. \qquad (2)$$

Unlike Equation (1) that determines the position of a module alone, Equation (2) contains multiple variables for both x and y coordinates for modules in a cluster. The work [13] further extends the median idea to determining the location of a seed module in a cluster and then placing the remaining modules accordingly, based on their relative positions to the seed module.

2.3.2 *Applications of Generalized Force-Directed Relaxation*

The work [1] presents a multilevel optimization for large-scale circuit placement, where a placement region is partitioned into a set of regular bins. At each level of refinement, this work locally permutes modules in a small subset of bins to improve the total wirelength. This work revises the ϵ-neighborhood, λ-exchange procedure as follows. Starting from a module m, the revised scheme computes its ϵ-neighbors and randomly selects one module, and for this newly selected module, this scheme further computes its ϵ-neighbors; this process is repeated until λ modules are selected. For these λ modules, the scheme tries all placement permutations and selects the permutation with the smallest wirelngth for real exchange. For example, if modules A, B, J are selected, all six permutations will be considered: no exchange, $A \leftrightarrow B$, $A \leftrightarrow J$, $B \leftrightarrow J$, $A \rightarrow B \rightarrow J \rightarrow A$, $A \rightarrow J \rightarrow B \rightarrow A$. The authors claimed to find superior solutions to those by the original work [6]. The POLAR placer also extends the GFDR framework in its density preserving refinement process [9].

3. FUTURE DIRECTIONS

Although placement is a classical problem, modern design challenges have reshaped the problem. The modern placement problem becomes very complicated mainly because of many emerging challenges with the following four aspects: (1) design scalability: handle ultra large-scale designs for modern applications; (2) multi-objective requirements: consider multiple placement constraints simultaneously, such as preplaced blockages, routability, timing, reli-

ability, co-design with other circuit components (clock networks, power/ground networks), etc.; (3) heterogeneous circuit components: tackle standard cells of different heights, mixed-sized designs with thousands of big macros together with tens of millions of standard cells, heterogeneous circuit components in an FPGA, etc.; and (4) emerging technologies: handle 3D placement, discrete FinFET-based placement, manufacturability-aware placement, etc. As a result, modern placement problems have attracted much research attention recently. We believe that the methods of the OR and GFDR will still play an important role in modern and future placement problems. In the following subsections, we present some potential research directions for modern placement with these challenges.

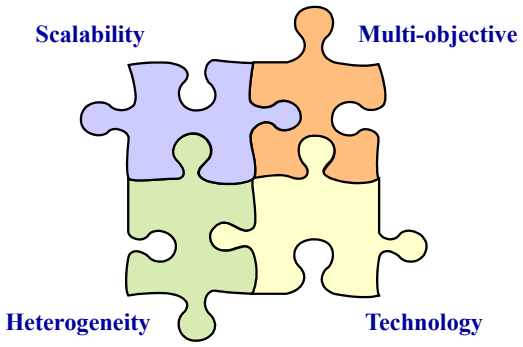

Figure 5: Major current and future circuit placement challenges.

3.1 Multi-Cell-Height Placement

For modern circuit designs, standard-cell libraries might contain cells of different heights, say, single-row-height cells, double-row-height cells, etc. Higher cells provide greater drive strengths at the cost of larger areas and power. Such mixed-cell-height cells incur complicated challenges for placement, because of the heterogeneity in cell dimensions [5, 10, 14]. As illustrated in Figure 6, mixed-cell-height placement shall consider standard cells of different cell heights and power-rail alignment as well. For an odd-row-height cell, such alignment can be achieved also by vertical cell flipping, while there are two types of an even-row-height cell with either VDD or VSS running along its top and bottom boundaries. The mixed-cell-height placement incurs new challenges for the computation of the ORs and thus force-directed relaxation, and its bin selection from ϵ-neighbors as well, especially when additional design constraints (e.g., the minimum implant area) need to be addressed simultaneously.

3.2 Mixed-Size Placement

A modern chip could contains thousands of big macros (due to IP modules, embedded memory modules, analog modules, etc.) and tens of millions of small standard cells, which significantly differ in both sizes and shapes. Figure 7

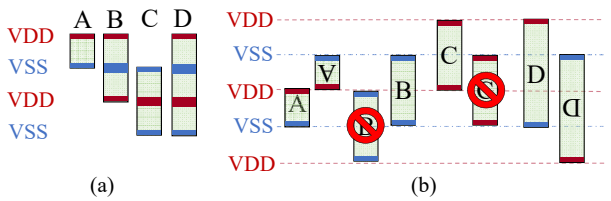

Figure 6: Mixed-cell-height placement shall consider standard cells of different heights and power-rail alignment.

shows two instances of mixed-size placements with large macros, with a single macro hierarchy (single-domain mixed-size placement in Figure 7(a)) and multiple macro hierarchies (multi-domain mixed-size placement in Figure 7(b)) with region constraints.

Pre-designed macros typically preserve multiple metal layers for interior routing, and these regions could become blockages during routing. Consequently, macros have a significant impact on chip routability. Further, the optimal regions and thus force-directed relaxation would be significantly different from a design with standard cells alone. So a modern placer should be capable of handling macro orientations and positions and capturing the interactions between big macros and small standard cells to derive accurate models for the optimal region computation and ϵ-neighbor selection for placement optimization.

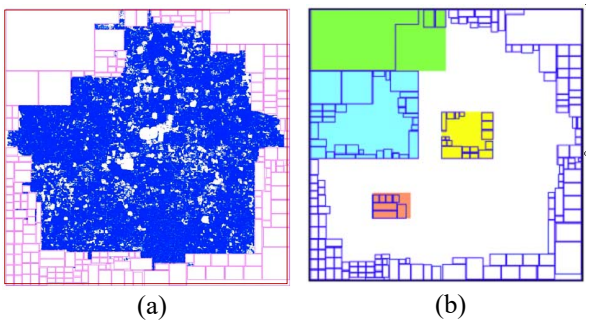

Figure 7: Mixed-size placements with large macros. (a) Single-domain mixed-size placement. (b) Multi-domain mixed-size placement with region constraints.

3.3 Routability/Timing-Driven Placement

Traditional placement relies on total wirelength minimization to obtain better circuit performance and smaller layout area. However, there is a mismatch between wirelength and congestion objectives in placement. See Figure 8 for the dramatically different behaviors with wirelength- and routability-driven placements. A wirelength-driven placer (NTUplace3 [3]) packs standard cells closer to minimize the total wirelength (see Figure 8(a)), incurring significant routing congestion violations (see Figure 8(b)). In contrast, a routability-driven placer (NTUplace4 [7]) spreads standard

cells over the chip to achieve better routability at the cost of longer total wirelength (Figures 8(c) and (d)). The original OR and ϵ-neighborhood formulations are intended for wirelength optimization alone. It is thus desirable to develop a routabilty-driven OR formulation and its corresponding force-directed relaxation scheme to optimize routabilty and wirelength simultaneously.

Timing optimization during placement is critical to high-speed circuit designs. Traditional placement algorithms often try to achieve the timing goal via wirelength minimization. Nevertheless, there is a gap between wirelength and actual delay, so many methods have been proposed to overcome this challenge. Existing timing-driven placement algorithms can be classified into two major categories: (1) path-based and (2) net-based methods [2]. Net-based methods are much more popular because the prohibitive exponentially-growing number of timing paths for the path-based methods. The net-based method converts the timing constraint of each path into net weights. For a placement algorithm with the OR and ϵ-neighborhood formulation, such net weight modeling is crucial for developing a timing optimization technique with high accuracy, low complexity, and good controllability.

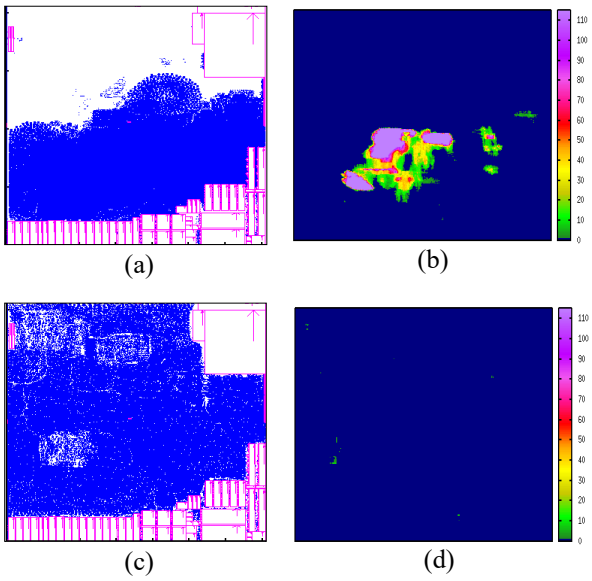

Figure 8: Dramatically different behaviors with wirelength- and routability-driven placements on the circuit sb12. (a) Wirelength-driven placed layout. (b) Congestion map for the wirelength-driven placed layout. (c) Routability-driven placed layout. (d) Congestion map for the routability-driven placed layout.

3.4 FinFET Self-Heating-Aware Placement

With their lower threshold voltage and smaller dynamic current leakage than those of traditional planar devices, FinFETs have emerged as a popular 3D transistor technology for circuit designs at the 22nm node and beyond [8]. The number of fins in a FinFET transistor plays a key role in determining its circuit performance and the self-heating effect caused by fingers is getting more severe, due mainly to the low thermal conductivity of buried oxide and interlayer dielectric materials and its compacted 3D device geometry. As a result, the self heating could significantly cause performance and reliability degradation. The self-heating effect is more dominant between fins and fins than that between devices and devices [16, 17]; see Figure 9 for an illustration. So a device itself acts like a thermal source. It is thus desirable to consider placement of such thermal sources to reduce the self-heating effect for designs with the FinFET technology. To handle the self-heating-aware placement problem, it is of particular importance to develop an effective model of the thermally optimal region and incorporate such a model into an effective placement framework to achieve desired solution quality.

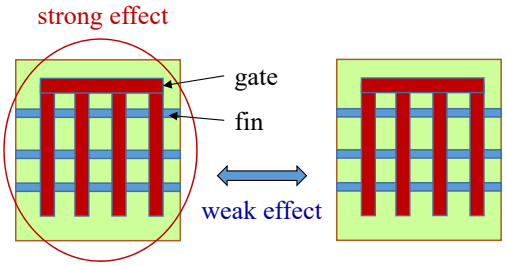

Figure 9: Self-heating effect inside a FinFET device and between two FinFET devices. The effect is more dominant between fins and fins than that between devices and devices.

3.5 FPGA Placement

A traditional symmetrical-array-based FPGA contains a two-dimensional array of configurable logic blocks (CLBs) surrounded by general routing resources and bounded by I/O blocks [4]. A modern FPGA often consists of complex heterogenous blocks, such as RAMs and DSPs, widely used to implement various circuit applications effectively. These complex blocks often contain datapath-intensive circuits. It is desirable to develop novel techniques to handle large-scale heterogeneous FPGAs placement with issues on heterogeneity, datapath regularity, and scalability. Obviously, any methods involving the ORs and force-directed relaxation would need to be revised to address the unique problems induced from the heterogeneous FPGA structures.

4. CONCLUSIONS

This article has introduced popular algorithmic paradigms for circuit placement. The classical, yet effective GFDR method and the OR formulation for finding desired placement solutions have then been presented. We have also discussed their impacts on modern placement and applications. Finally, we have further provided future placement research directions associated with the OR and GFDR formulations.

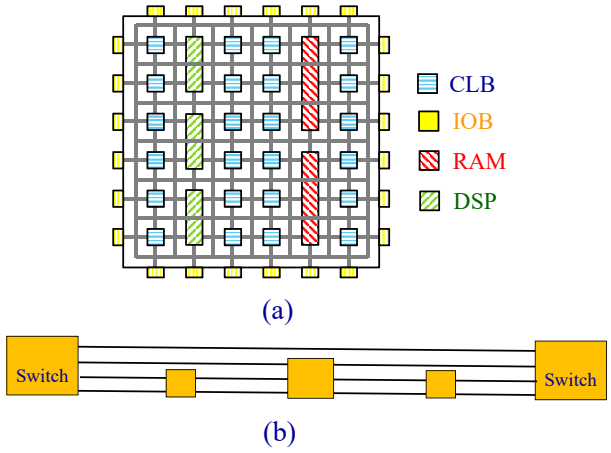

Figure 10: An FPGA architecture with heterogeneous circuit components (CLBs, IOBs, RAMs, and DSPs) and segmented routing structures.

A significant role of the OR and GFDR formulations for future placement problems is well expected.

5. REFERENCES

[1] T. F. Chan, J. Cong, T. Kong, and J. R. Shinnerl, "Multilevel optimization for large-scale circuit placement," in *Proc. IEEE/ACM Int. Conf. Computer-Aided Design*, pp. 171–176, San Jose, CA, November 2005.

[2] Y.-W. Chang, Z.-W. Jiang, and T.-C. Chen, "Essential issues in analytical placement algorithms," *IPSJ Transactions on System LSI Design Methodology*, pp. 815–835, August 2009

[3] T.-C. Chen, Z.-W. Jiang, T.-C. Hsu, H.-C. Chen, and Y.-W. Chang, "NTUplace3: An analytical placer for large-scale mixed-size designs with preplaced blocks and density constraints," *IEEE Trans. Computer-Aided Design of Integrated Circuits and Systems*, Vol. 27, No. 7, pp. 1228–1240, July 2008.

[4] Y.-C. Chen, S.-Y. Chen, and Y.-W. Chang, "Efficient and effective packing and analytical placement for large-scale heterogeneous FPGAs," in *Proc. IEEE/ACM Int. Conf. Computer-Aided Design*, San Jose, CA, November 2014.

[5] W.-K. Chow, C-.W. Pui, F.Y. Young, "Legalization algorithm for multiple-row height standard cell design," in *Proc. ACM/IEEE Design Automation Conference*, Austin, TX, June 2016.

[6] S. Goto, "An efficient algorithm for the two-dimensional placemet problem in electrical circuit layout," *IEEE Trans. Circuits and Systems*, Vol. 28, No. 1, pp. 12–18, January 1981.

[7] M.-K. Hsu, Y.-F. Chen, C.-C. Huang, S. Chou, T.-H. Lin, T.-C. Chen, and Y.-W. Chang, "NTUplace4h: A novel routability-driven placement algorithm for hierarchical mixed-size circuit designs," *IEEE Trans. Computer-Aided Design of Integrated Circuits and Systems*, Vol. 33, No. 12, pp. 1914ąX1927, December 2014.

[8] D. Hisamoto, W.-C. Lee, J. Kedzierski, H. Takeuchi, K. Asano, C. Kuo, E. AAnderson, T.-J. King, J. Bokor, and C. Hu, "FinFET-a self-aligned double-gate MOSFET scalable to 20 nm," *IEEE Trans. Electron Devices*, vol. 47, No. 12, pp. 2320ąV-2325, December 2000.

[9] T. Lin, C. Chu, J. R. Shinnerl, I. Bustany, and I. Nedelchev, "POLAR: Placement based on novel rough legalization and refinement," in *Proc. IEEE/ACM Int. Conf. Computer-Aided Design*, San Jose, CA, November 2013.

[10] Y. Lin, B. Yu, X. Xu, J.-R. Gao, N. Viswanathan, W.-H. Liu, Z. Li, C. J. Alpert, and D. Z. Pan, "MrDP: Multiple-row detailed placement of heterogeneous-sized cells for advanced nodes," in *Proc. IEEE/ACM Int. Conf. Computer-Aided Design*, Austin, TX, Novvember 2016

[11] M. Pan, N. Viswanathan, and C. Chu, "An efficient and effective detailed placement algorithm," in *Proc. IEEE/ACM Int. Conf. Computer-Aided Design*, pp. 48–55, San Jose, CA, November 2005.

[12] S. M. Sait and H. Youssef, *VLSI Physical Design Automation: Theory and Practice*, World Scientific Publishing Co., 1999

[13] K.-H. Tseng, Y.-W. Chang, and C. C. C. Liu, "Minimum-implant-area-aware detailed placement with spacing constraints," in *Proc. ACM/IEEE Design Automation Conf.*, Austin, TX, June 2016.

[14] C.-H. Wang, Y.-Y. Wu, J. Chen, Y.-W. Chang, S.-Y. Kuo, W. Zhu, and G. Fan, "An effective legalization algorithm for mixed-cell-height standard cells," in *Proc. of IEEE/ACM Asia and South Pacific Design Automation Conf.*, January 2017.

[15] L. T. Wang, Y.-W. Chang, and K.-T. Cheng (Ed.), *Electronic Design Automation: Synthesis, Verification, and Test*, Morgan Kaufmann, 2009

[16] C. Xu, S. K. Kolluri, K. Endo, and K. Banerjee, "Analytical thermal model for self-heating in advanced FinFET devices with implications for design and reliability," *IEEE Trans. Computer-Aided Design of Integrated Circuits and Systems*, pp. 1045-1058, Vol. 32, No. 7, July 2013.

[17] J.-Y. Yan, Y.-J. Peng, and C. W. Liu, "TCAD simulation and analytic modeling of self-heating effects in FinFETs," *MediaTek Semi-annual Report*, June 2016.