International Journal of Scientific Research & Engineering Trends
Volume 9, Issue 3, May-Jun-2023, ISSN (Online): 2395-566X

Artificial Intelligence in VLSI Routing: A Review

Nagavalli S, Jyoti R Munavalli
Dept of EC
BNM College of Engineering,
Bengaluru, Karnataka

Abstract- An Artificial Intelligence Approach to VLSI Routing presents a system that performs routing close to what human designers do. This paper summarises the different algorithms of AI/ML used in VLSI routing. This paper reviews on the different AI/ML techniques and algorithms to overcome routing problems like congestion prediction, to solve fundamental issues of routing like scalability, reward design, and end-to end learning paradigm, circuit routing, artificial intelligence enabled routing (AIER) mechanism with congestion avoidance in Software development networks.

Keywords- Artificial Intelligence, VLSI, routing, Software development networks.

I. INTRODUCTION

An evident challenge ahead for the integrated circuit (IC) industry in the nanometre regime is the investigation and development of methods that can reduce the design complexity ensuing from growing process variations and curtail the turnaround time of chip manufacturing. Conventional methodologies employed for such tasks are largely manual; thus, time-consuming and resource intensive.

In contrast, the unique learning strategies of artificial intelligence (AI) provide numerous exciting, automated approaches for handling complex and data intensive tasks in very-large-scale integration (VLSI) design and testing. Employing AI and machine learning (ML) algorithms in VLSI design and manufacturing reduces the time and effort for understanding and processing the data within and across different abstraction levels via automated learning algorithms. It, in turn, improves the IC yield and reduces the manufacturing turnaround time.

II. BASICS OF ROUTING

1. Types of Routing:

Routing is the process of laying physical connections to circuit blocks and pins assigned during placement. Global routing (GR) partitions the routing region into tiles and decides tile-to-tile paths for all nets while attempting to optimize objectives such as minimum wire length and timing budget. The actual geometric layout of each net within the assigned routing regions is carried out in the detailed routing stage. Routing congestion is the major bottleneck in the GR stage and DRVs are another area for improvement. Machine-learning-based approaches have been used to address these challenges. This paper reports on the different algorithms or techniques of ML/AI in VLSI Routing. Routing is typically a very complex combinatorial problem and is usually solved by using a two-stage approach of global routing followed by detailed routing.

Global routing partitions the routing region into tiles and decides tile-to-tile paths for all nets, while detailed routing assigns actual tracks and vias for nets. Fig. 1 illustrates the process of global routing and detailed routing. Global routing divides the routing region into tiles as shown in Fig (a) and generates a “loose” route for each connection by finding the tile-to-tile paths to connect pins and/or pads as shown in Fig (b). Detailing routing determines the exact route for each net by searching within the tile-to-tile path as shown in Fig(c).

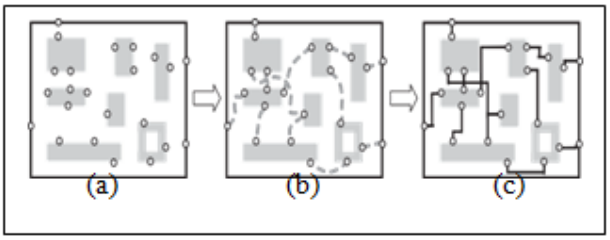

Fig1. (a) Global routing divides the routing region into tiles.(b) generates a “loose” route for each connection by finding the tile-to- tile paths to connect pins and/or pads.(c) Detailing routing determines the exact route for each net by searching within the tile- to-tile path.

2. Routing Model:

Applying the graph-search technique for routing requires modelling the routing resource as a graph where the graph topology can represent the chip structure. In graph modelling, the chip (routing region) is first partitioned into an array of rectangular tiles (or called global-routing tiles), each of which may accommodate tens of routing tracks in each dimension. A node in the routing graph represents a tile in the chip, while an edge denotes the boundary between two adjacent tiles. Each edge is assigned a capacity according to the physical routing area or the number of tracks in a tile. For detailed routing, the router decides the physical interconnections of nets by allocating wires on each metal layer and vias for switching between metal layers. There are two different

layer models, the reserved and unreserved, which allow the placement of wires with any directions. There are two kinds of detailed-routing models: the grid-based and grid less models as shown in Fig 2.

Grid-based routing involves superimposing a routing grid on the routing region, while grid less routing involves finding routing paths in the grid. The space between adjacent grid lines is called wire pitch, which is larger than or equal to the sum of the minimum width and spacing of wires. Grid-based detailed routing is much more efficient and easier for implementation. Grid less detailed routing is any model that does not follow the grid-based model and can use different wire widths and spacing. Various grids less models have been proposed, such as the connection graph, the implicit connection graph, the implicit triple-line graph, and corner stitching.

Grids less routing has greater flexibility and is more suitable for interconnect tuning optimization, such as wire sizing and perturbation. However, it is generally much slower than the grid-based one due to its higher complexity.

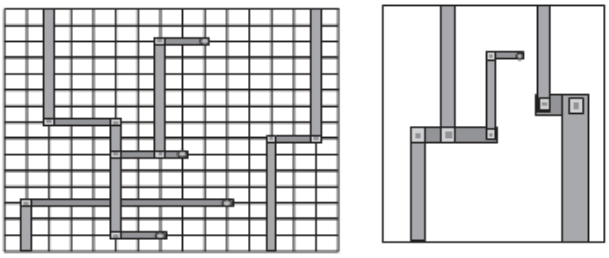

Fig 2. (a) Grid-based detailed routing.(b) Gridless detailed routing.

III. RELATED WORK

This section presents previous work done on the application of AI in VLSI routing protocols.

Ruoyu Cheng Junchi Yan [1] infers that Machine learning is an emerging tool for solving placement and routing problems, and DeepPlace and DeepPR are joint learning methods to achieve end-to-end placement and routing. Learning-based methods for placement, especially reinforcement learning (RL), have been proposed to obtain the generalization ability. Classical and learning-based methods for routing have also been proposed, such as rip-up and reroute, force-directed routing, and region-wise routing. This work combines reinforcement learning with a gradient based optimization scheme without clustering to obtain the full placement solution through end-to-end learning scheme. However, their model is limited to a moderate number of macros for placement due to the scalability issue of GNN and the large action space for reinforcement learning.

Y. V. Sneha et al. [2] worked on the prediction of network congestion at router using ML technique. When a burst of packets enters a network, the existing capacity may not be enough to support the traffic, leading to congestion. To avoid packet loss, effective measures must be taken to reduce the packet generation rate at source. Existing protocols are predefined mappings between the observed state and the corresponding action but are unable to adapt to the new environment or learn from past experience. To overcome these issues, the Machine Learning (ML) technique is needed to learn from experience and analyse the current network scenario to take certain actions. The training dataset consists of queue parameters such as current number of packets stored in queue (curpkts), current queue percentage (cur%), remaining packets that can be stored in queue (rempkts), remaining queue percentage (rem%), current queue average (curavg), old average (oldavg) and class label indicating whether there occurs congestion or not.

Simulation is performed to generate a dataset of predicting attributes. The most important details in this text are the steps of model construction, validation, deployment, and inference. A simple topology is chosen for simulation, consisting of a single router and a queue with a maximum size of 50 packets. Performance metrics are used to compare the performance of the model, with the Naive Bayes method being better than the SVM method.

Yan-Jing Wu et al. [3] experimented on the enabled routing in software defined networking where artificial intelligence can be implemented. Software defined networking (SDN) is an emerging networking architecture that separates the control plane from the data plane and moves network management to a central point. This paper presents an artificial intelligence enabled routing (AIER) mechanism with congestion avoidance in SDN, which can alleviate the impact of monitoring periods with dynamic routing and provide learning ability and superior route decisions. Performance metrics such as average throughput, packet loss ratio, and packet delay versus data rate are demonstrated. An OpenFlow switch comprises one or more flow tables, a group table, and an OpenFlow channel for the external controller. The BPA algorithm is used to train the ANN model.

The training data set is divided into a training data subset and a test data subset, with the training data subset used to train the ANN model and the test data subset used to verify the accuracy of the trained ANN model. The ANN model is used to calculate the congestion probability for each path configuration, and the AIER mechanism can monitor the queuing length of each OpenFlow switch to avoid inaccurate output in the trained ANN model. Performance analysis uses the Dijkstra algorithm, OpenFlow Protocol V1.3, and the Iperf tool to generate UDP flows at data rates ranging from 70 Mbps to 150

Mbps. The bandwidth of each link is 250 Mbps, and the size of each OpenFlow switch is 200 packets. The monitoring period is fixed at 3, 5, or 10 s, and a multilayer perceptron (MLP) is used as an ANN model in the control plane. 65,000 data records were collected to train the ANN model, with 120 and 140 neutrons at the first and second hidden layers being adopted. Performance measures include average throughput, packet loss ratio, and packet delay for different monitoring periods. The AIER mechanism is superior to static routing due to its ability to change transmission paths periodically if a predefined congestion condition occurs. Dynamic routing can adjust path allocation periodically to alleviate congestion in the shortest path, resulting in a smaller packet loss ratio compared with static routing. The AIER mechanism demonstrates a significant improvement in packet loss owing to its intelligent path selection design with congestion avoidance.

A thesis in the University of British Columbia [4] explains about Machine Learning Based Techniques for Routing Interconnects in Very Large Scale Integrated (VLSI) Circuits. The congestion estimation algorithms found in routers tend to be probabilistic-based or global-routing- based. However, ML algorithms can be trained to accurately estimate the congestion of new circuits based on example designs. The MARS (Multivariate Adaptive Regression Splin) approach can construct a congestion estimation model which considers different information within each routing graph vertex, such as the number and locations of local pins and nets of each vertex. The key challenge in applying ML to EDA applications is how to map a circuit into a format readily "understandable" by ML algorithms chosen. The MARS algorithm is a "matrix-vector pair" format for the design data in the netlist. The ISPD'08 benchmarks were used to experiment the congestion estimation quality.

The experiments were run on a Linux workstation with 4-core 2.6GHz CPU and 8GB memory. The 2-pin path availability feature in the MARS was set to consider L - shapes, and the difference with and without MARS predictive model was measured in terms of runtime, initial routing solutions, and final routing solutions. The results showed that the predictive model was 94% faster than conventional congestion estimation methods when applied on the ISPD'08 benchmark suite, and the router achieved an overall 4.8% runtime speedup in global routing using MARS predictive model. This can significantly benefit current industry designs.

Youbiao He et al. [5] proposes a general routing approach that solves circuit routing by Monte Carlo tree search (MCTS) with deep neural network (DNN) guided rollout. Experiments on randomly generated single-layer circuits show the potential to route complex circuits. The proposed approach can solve problems that benchmark methods cannot solve and outperform the vanilla MCTS approach. MCTS-based circuit routing is a search algorithm for deciding optimal actions to take at discrete steps, used to construct a path from one of the two pins of a net, expanding it by one vertex each step. Figure 4 shows an example of routing in our approach and traditional sequential routing. Traditional routing approaches connect nets 1, 2, 4 and 5 with shortest paths, while our approach can leave space for nets to be routed later.

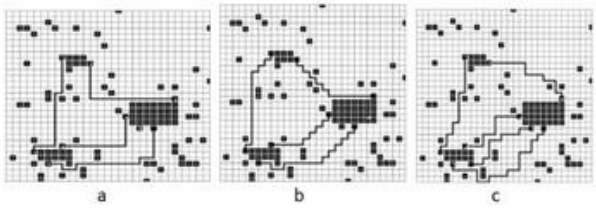

Fig 3. (a) 100 iterations (b) 500 iterations (c) 2000 iterations

Tingyuan Niea1 et al. [6] proposes a novel complex network machine learn (CNML) framework for predicting routing congestion in modern FPGA designs during the placement stage. It uses an artificial neural network (ANN) model to achieve outstanding performance with the second-best rank in accuracy and inference runtime. The proposed methodology utilizes the widely used academic placer UTPlaceF to derive the placement, which includes cells and wires that form a graph with connectivity described in the netlist. To create the undirected weighted network, we treat modules in the circuit placement as nodes, interconnections between two modules as edges, and the Manhattan Distance between two modules as the edge weight. This text discusses a novel approach to predict routing congestion of a circuit's physical design by transforming the design into a complex network and extracting features to estimate the routing congestion by machine learning models.

The proposed framework outperforms several congestion estimation techniques in terms of accuracy metrics such as R2, MAE, and RMSE. The average inference runtime consumes 22 milliseconds, which shows high effectiveness for the prediction. The importance of complex network features in CNML learning was evaluated and showed better generalization capability. In the future, the CNML will be integrated into EDA tools to guide subsequent optimization.

IV. CONCLUSION

Employing AI and machine learning (ML) algorithms in VLSI design and manufacturing reduces the time and effort for understanding and processing the data within and across different abstraction levels. It, in turn, improves the IC yield and reduces the manufacturing turnaround time. The market for hardware ML/AI accelerators is growing rapidly, with most big

semiconductor IP/SoC names (Intel, Samsung, ARM …) producing their own ML accelerators. Knowledge of ML is beneficial even if you are just reusing an existing algorithm. VLSI verification is also expected to benefit from ML. ML models can be trained to evaluate the effectiveness of test-benches in coverage-driven verification. Constrained random verification can be guided using ML models for more effective coverage.

V. ACKNOWLEDGEMENT

The authors would like to thank B N M Institute of Technology, Bangalore for all the support and encouragement.

REFERENCES

[1] "On Joint Learning for Solving Placement and Routing in Chip Design(2021)" Ruoyu Cheng Junchi Yan, Department of Computer Science and Engineering MoE Key Lab of Artificial Intelligence, AI Institute Shanghai Jiao Tong University, Shanghai, China.

[2] "Prediction of Network Congestion at Router using Machine learning Technique(2020)" Y. V. Sneha, Vimitha, Vishwasini, Shravan Boloor, Adesh N. D., Dept. Computer Science and Engineering, Shri Madhwa Vadiraja Institute of Technology and Management, Bantakal, Udupi, India, Dept. Information and Communication Technology, Manipal Institute of Technology and Manipal Academy of Higher Education, Manipal, Udupi, India

[3] "Artificial Intelligence Enabled Routing in Software Defined Networking"(article)(2020) Yan-Jing Wu 1,Po-Chun Hwang 2 , Wen-Shyang Hwang 2 and Ming-Hua Cheng

[4] "Machine Learning Based Techniques for Routing Interconnects in Very Large Scale Integrated (VLSI) Circuits(2022)", a thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the faculty of graduate and postdoctoral studies (Electrical and Computer Engineering) The University of British Columbia.

[5] "Circuit Routing Using Monte Carlo Tree Search and Deep Neural Networks" by Youbiao He and Forrest Sheng Bao Dept. of Computer Science, Iowa State University, Ames, IA 50014(2022)

[6] "Complex Network Machine Learning for FPGA Routing 2 Congestion Prediction"(2023)by Tingyuan Niea1, Qi Konga, Pengfei Liua, Zhenhao Wanga, Kun Zhaoa School of Information and Control Engineering, Qingdao University of Technology, No. 5 777, Jialingjiang Road, Qingdao, 266052/Shandong, China.