DeepTH: Chip Placement with Deep Reinforcement Learning Using a Three-Head Policy Network

Dengwei Zhao
Dept. of Computer Science and Engineering
Shanghai Jiao Tong University
Shanghai, China

Shuai Yuan
Dept. of Micro-Nano Electronics
Shanghai Jiao Tong University
Shanghai, China

Yanan Sun*
Dept. of Micro-Nano Electronics
Shanghai Jiao Tong University
Shanghai, China

Shikui Tu*
Dept. of Computer Science and Engineering
Shanghai Jiao Tong University
Shanghai, China

Lei Xu
Dept. of Computer Science and Engineering
Shanghai Jiao Tong University
Shanghai, China

***Abstract*—Modern very-large-scale integrated (VLSI) circuit placement with huge state space is a critical task for achieving layouts with high performance. Recently, reinforcement learning (RL) algorithms have made a promising breakthrough to dramatically save design time than human effort. However, the previous RL-based works either require a large dataset of chip placements for pre-training or produce illegal final placement solutions. In this paper, DeepTH, a three-head policy gradient placer, is proposed to learn from scratch without the need of pre-training, and generate superior chip floorplans. Graph neural network is initially adopted to extract the features from nodes and nets of chips for estimating the policy and value. To efficiently improve the quality of floorplans, a reconstruction head is employed in the RL network to recover the visual representation of the current placement, by enriching the extracted features of placement embedding. Besides, the reconstruction error is used as a bonus during training to encourage exploration while alleviating the sparse reward problem. Furthermore, the expert knowledge of floorplanning preference is embedded into the decision process to narrow down the potential action space. Experiment results on the ISPD 2005 benchmark have shown that our method achieves 19.02% HPWL improvement than the analytic placer DREAMPlace and 19.89% improvement at least than the state-of-the-art RL algorithms.**

***Index Terms*—Chip placement, Reinforcement learning, Intrinsic reward, Three head, Visual reconstruction.**

I. Introduction

Placement is one of the most complex and time-consuming stages in the chip physical design process. Macros and standard cells are placed on the chip canvas to optimize the power consumption, performance, and area (PPA), while meeting constraints on density and routing congestion. With the continuous growth in the scale and design complexity of modern VLSI circuits, the placement algorithms need to quickly solve the increasingly complex multi-objective optimization problems involving a large number of iterations.

This work was supported by National Natural Science Foundation of China (NSFC) under Grant No. 62174110, National Key RD Program of China under Grant No. 2021YFA0717400, and BirenTech Research. (* *Corresponding authors: Yanan Sun and Shikui Tu. Emails: sunyanan@sjtu.edu.cn, tushikui@sjtu.edu.cn*)

Recently, reinforcement learning (RL) algorithms [2] & [3] have made a promising breakthrough to dramatically save design time than human effort. However, RL algorithms typically suffer from sparse rewards due to the huge space complexity of the chip placement problem. In this paper, DeepTH, a **T**hree-**H**ead policy gradient placer, is proposed to solve chip placement problem and our contributions are summarized from the following aspects.

- Based on the typical policy-value network, a reconstruction head for recovering another visual placement is added to help the graph network extract more reliable features.
- The reconstruction error is treated as an intrinsic reward to encourage exploration while placing chip objects, which alleviates the sparse reward problem effectively.
- Expert knowledge is embedded into the playing agent's decision process, screening unreliable actions to narrow down the potential state space.
- Placements provided by the previous state-of-the-art work [3] is illegal due to the existence of macro overlaps. Our algorithm adopts a legalization algorithm to eliminate the overlap, obtaining more meaningful placements. The average wirelength on ISPD 2005 benchmark achieves 19.72% HPWL improvement than [3]. The source code is available on https://github.com/CMACH508/DeepTH.

II. Method

In this paper, DeepTH is proposed to perform macro placement and standard cells are placed by an analytical algorithm. A hypergraph partitioning algorithm is used to first divide standard cells into disjoint clusters based on their netlist connections. Standard cell clusters and macros will be placed on the gridded canvas according to the policy in the order of area from largest to smallest sequentially. When the macros' location is determined, standard cells will be repositioned by DREAMPlace 3.0 [1], which also gives the evaluation of the final placement.

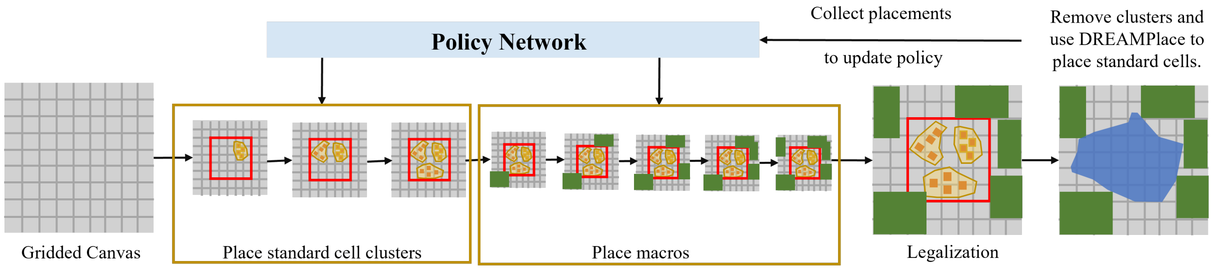

Fig. 1. Placement framework for our algorithm.

TABLE I
COMPARISON OF HPWL WITH DIFFERENT ALGORITHM ON ISPD 2005 BENCHMARK. (K IS THE NUMBER OF CHIPS BETTER THAN DREAMPLACE)

model	is legal	*adaptec*1	*adaptec*2	*adaptec*3	*adaptec*4	*bigblue*1	*bigblue*3	K	mean
DREAMPlace 3.0 [1] *	✓	128927038	152699768	175509798	281010687	103799877	426878464	0	100.00%
Google's placer [2] **	✓	86675688	124125088	257956784	255770672	168456768	460581568	3	109.45%
DeepPlace [3]	×	80117232	123265964	241072304	236391936	140435296	450633360	3	100.87%
DeepTH (ours)	✓	84905888	136776464	228162688	251242096	101607936	481772992	4	97.60%
DeepTH (fine-tune, ours)	✓	95768912	132401504	142752416	134953008	136484800	273440000	**5**	**80.98%**

* The HPWL of DREAMPlace is the mean of 20 execution results. ** The results of Google's placer are reproduced in DeepPlace [3].

A. Three-head Network Architecture

In general, neural networks in RL only have two heads to estimate the policy and value. A three-head network is proposed by employing a reconstruction head to recover the placement's visual representation. Object features and the adjacent matrix is provided as the network's input and each object is described by its width, height, whether is placed or not and the placed coordinates. Obviously, the input object features have enough information to reconstruct the current placement visually. Based on this observation, the global embedding is fed into the reconstruction head to recover the placement into a picture. Both real collected canvas B and reconstructed canvas $\hat{B}$ are further fed into function f, and the reconstruction error is

$$L_{Rec} = ||f(\hat{B}) - f(B)||^2. \quad (1)$$

B. Reward Function with Exploration bonus

Half-perimeter wire length $HPWL$ is adopted to evaluate the placement approximately. For the intermediate step, r_{wl} is always zero. Due to the huge state space of the chip placement problem, RL algorithm suffers from sparse rewards greatly. One common approach is to consider intrinsic rewards as exploration bonuses. Previous works propose that the agent's familiarity with the environment can be employed to measure curiosity. In this paper, canvas reconstruction error L_{Rec} is adopted as exploration bonus r_{eb}. Smaller L_{Rec} indicates that the reconstruction network is familiar with the current placement and less exploration reward is given. The final reward is the summation of HPWL and the intrinsic reward.

C. Expert Knowledge

Chip placement problem has been investigated for decades, and there exists a lot of expert experience, which can benefit the algorithm's learning. For example, it would be more appropriate to put macros in the marginal and reserve a continuous standard cell placement region in the center [4]. This observation is transformed as an extra constraint by masking marginal parts when placing standard cell clusters.

III. RESULT

PPO algorithm is adopted to train the policy network. Six of ISPD 2005 are used to evaluate the performance. Placer's relative performance is evaluated by

$$score = \frac{Placer\text{'s } HPWL}{DREAMPlace\text{'s } HPWL}. \quad (2)$$

The mean score of each placer in six testing chips is summarized in Table I and the placer with a smaller mean score performs better. DREAMPlace 3.0 [1] is employed to fine-tune our results to mitigate the effects of position changes during the legalization process. The fine-tune placement provided by our DeepTH achieves an average of 19.02% HPWL improvement than DREAMPlace, also 28.47% and 19.89% higher than Google's placer [2] and DeepPlace [3] respectively. Solutions provided by DeepTH are overlap-free while SOTA DeepPlace is not. Experiment results have demonstrated the superiority of our algorithm greatly.

REFERENCES

[1] J. Gu, Z. Jiang, Y. Lin, and D. Z. Pan, "Dreamplace 3.0: Multi-electrostatics based robust vlsi placement with region constraints," in *2020 IEEE/ACM International Conference On Computer Aided Design (ICCAD)*. IEEE, 2020, pp. 1–9.
[2] A. Mirhoseini, A. Goldie, M. Yazgan, J. W. Jiang, E. Songhori, S. Wang, Y.-J. Lee, E. Johnson, O. Pathak, A. Nazi *et al.*, "A graph placement methodology for fast chip design," *Nature*, vol. 594, no. 7862, pp. 207–212, 2021.
[3] R. Cheng and J. Yan, "On joint learning for solving placement and routing in chip design," *Advances in Neural Information Processing Systems*, vol. 34, pp. 16 508–16 519, 2021.
[4] Y. Liu, Z. Ju, Z. Li, M. Dong, H. Zhou, J. Wang, F. Yang, X. Zeng, and L. Shang, "Floorplanning with graph attention," in *Proceedings of the 59th ACM/IEEE Design Automation Conference*, 2022, pp. 1303–1308.