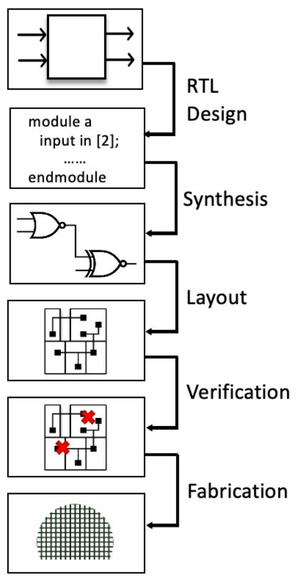
Intelligent Circuit Design and Implementation with Machine Learning in Design Automation

Zhiyao Xie
Advisor: Yiran Chen, Hai (Helen) Li
Duke University

Background and Takeaway



ML for EDA in commercial tools

cadence
Cadence Innovus™

SYNOPSYS®
Synopsys ICC™ II

.....

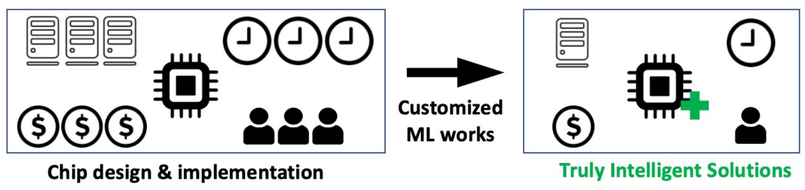
ML for Chip Design

Traditional Chip Design

Contributions of my Ph.D. works

- **Bridge** separated design steps with ML, covering RTL, netlist, layout, etc.
- Facilitate **early-stage** chip optimizations
- Not just ML for chip, also **ML in chip**
- Targeting **all** major chip design objectives

My vision

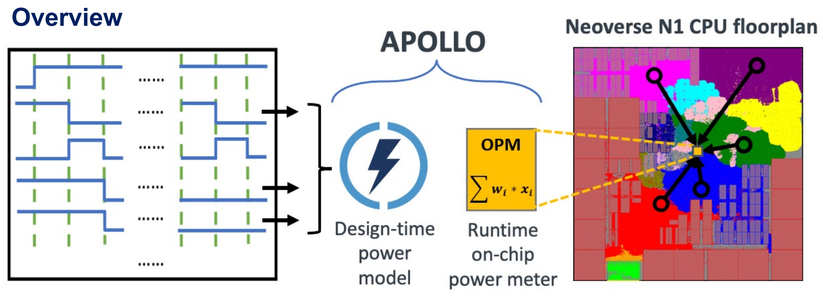


Power Modeling of RTL [MICRO'21] (Best Paper Award)

Highlights (my work APOLLO)

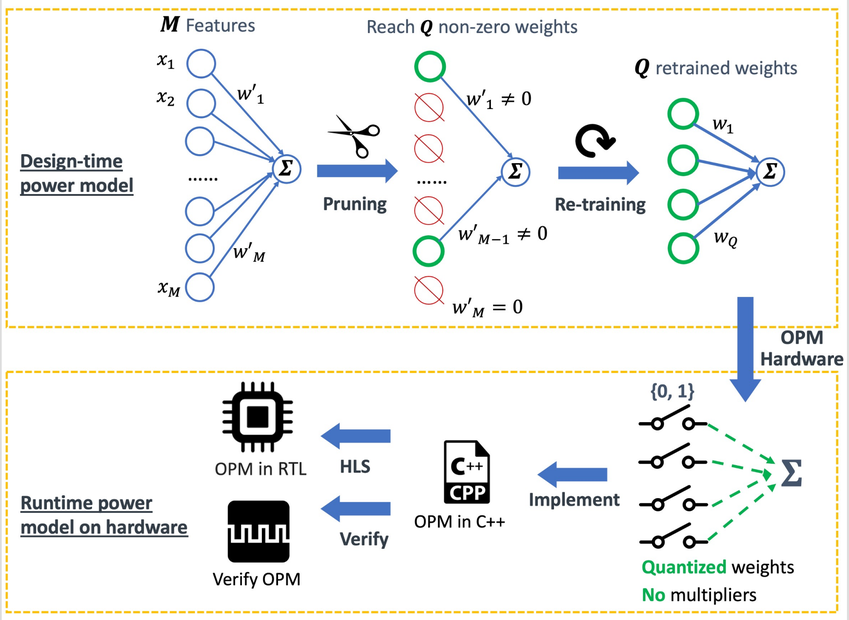
- **Fast** and **accurate** design-time power model handling millions-of-cycles benchmarks in minutes
- **Low-cost** and **accurate** runtime on-chip power meter (OPM)
- Unprecedented **per-cycle** temporal resolution
- Fully **automated** development process for any given design

Overview



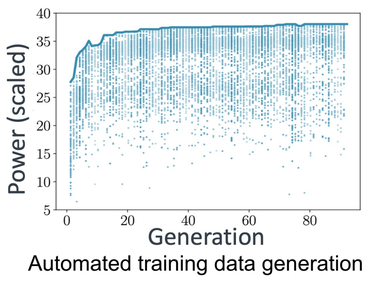
Key Method

- **Automatically** select only ~100 RTL signals as input power proxies
- The selection is performed by MCP-based feature pruning algorithm
- Construct OPM hardware **without** multipliers, with weight **quantization**



Experiment Setup

- Experiments on **high-volume commercial** microprocessors
- Training on **auto-generated** workloads, testing on Arm workloads



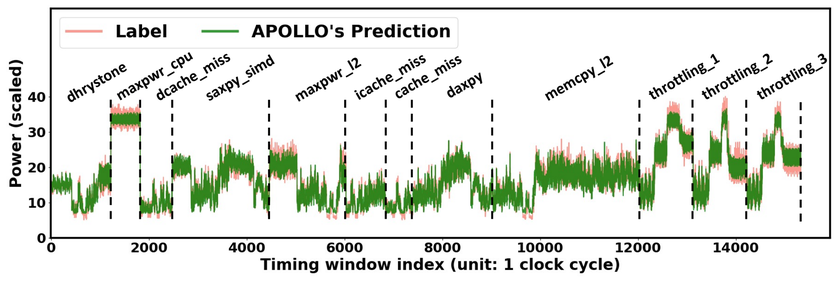
Automated training data generation

Neoverse N1 (infra)
In AWS Graviton

Cortex A77 (mobile)
In Snapdragon 865

Experimental Results

- **90-95%** accuracy of **17-million-cycles** SPEC workload in **minutes**
- **0.2% area overhead** for on-chip meter without accuracy degradation



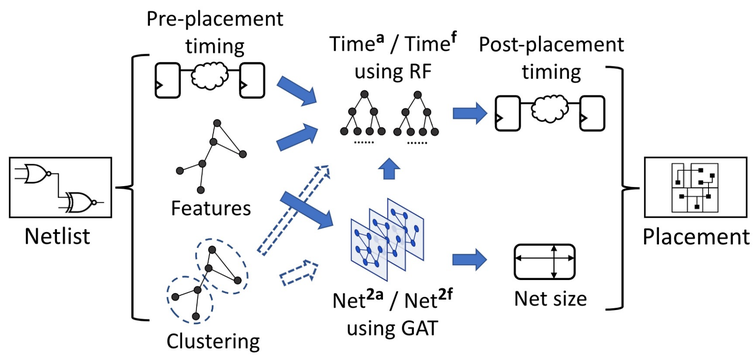
Interconnect of Netlist [ASP-DAC'21] [TCAD (Under-review)]

Highlights (my work Net²)

- **First** GNN-based method for pre-placement net length estimation
- **First** ML-based detailed timing estimator before placement

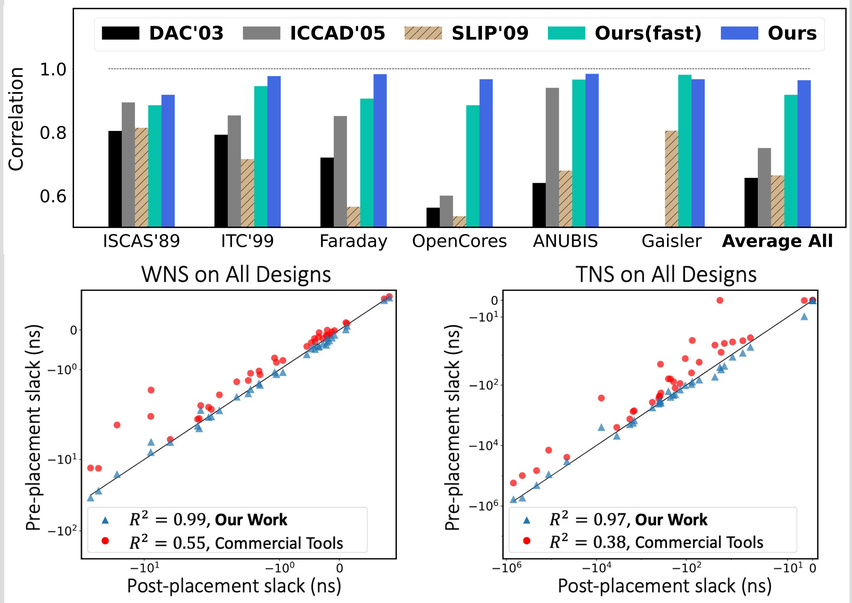
Key Method

- Extract global topology information through partitioning/clustering
- Customized graph attention network (GAT) method



Experimental Results

- **High accuracy** for individual net length prediction
- **Improve** slack estimations from commercial tools



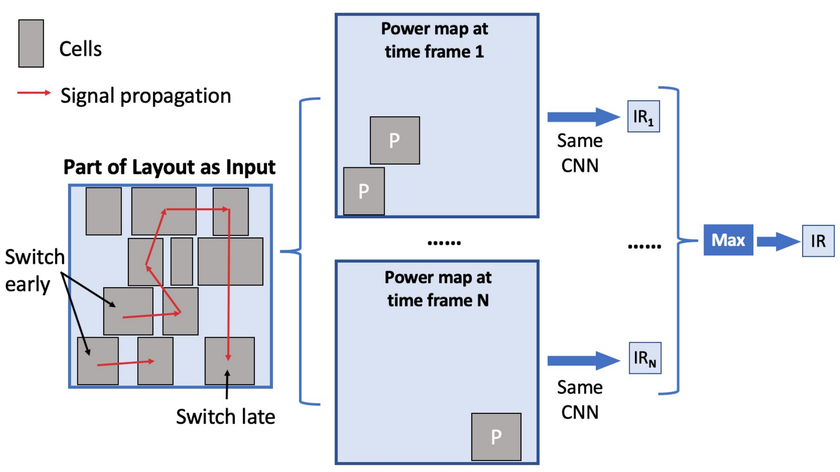
IR Drop Estimation of Layout [ASP-DAC'21] [ICCAD'20]

Highlights (my work PowerNet)

- **First** method that claims to perform **design-independent** fast IR drop estimations, for **both** vectorless and vector-based estimations
- **30X faster** than simulation-based commercial IR drop analysis tools

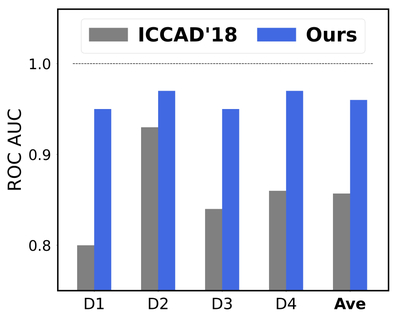
Key Method

- **Time-decomposed** power density as input features
- The **max** estimated IR drop among all time frames as final prediction



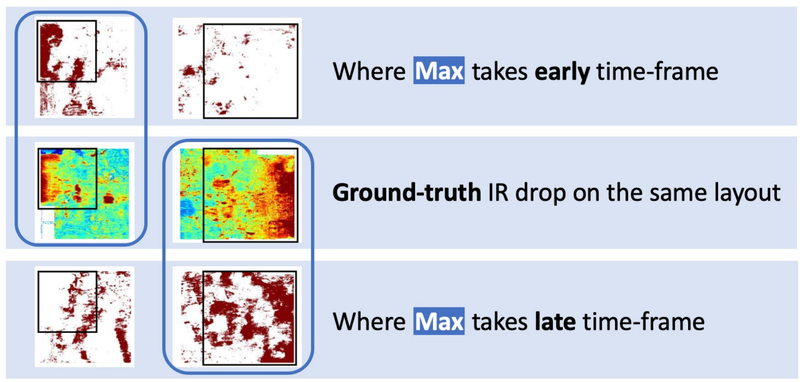
Experimental Results

- Integrated to guide mitigation flow to reduce IR violations by **20-30%**
- **Interpretability:** showing violations triggered at different time frames



Design MD1	Violated Cell	# Hotspots
Before Mitigate	22185	5092
After Mitigate	17052	3778
Improvement	23%	26%

Design MD2	Violated Cell	# Hotspots
Before Mitigate	31097	3627
After Mitigate	23941	2489
Improvement	23%	31%

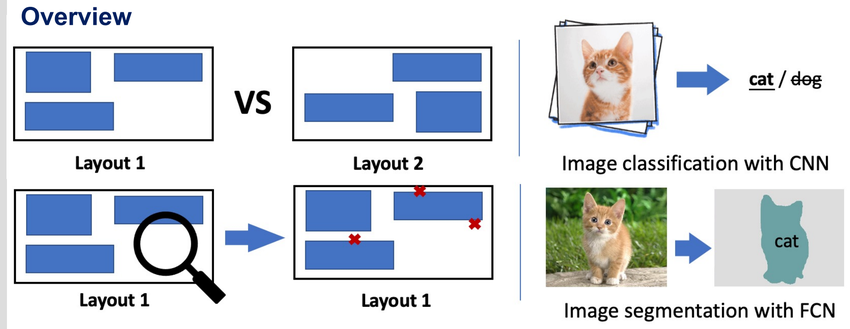


Routability Estimation of Layout [ICCAD'18]

Highlights (my work RouteNet)

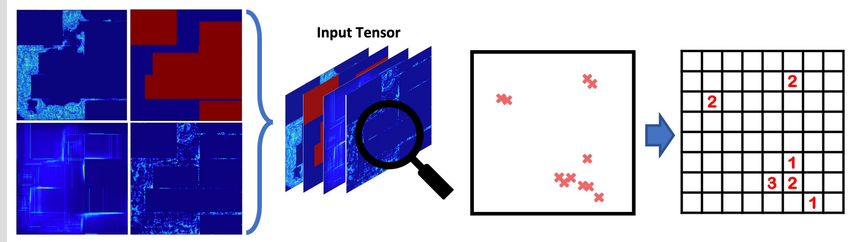
- **First** deep learning-based DRV estimator, capturing **global** information
- **Orders-of-magnitude speedup** compared with accurate simulations

Overview



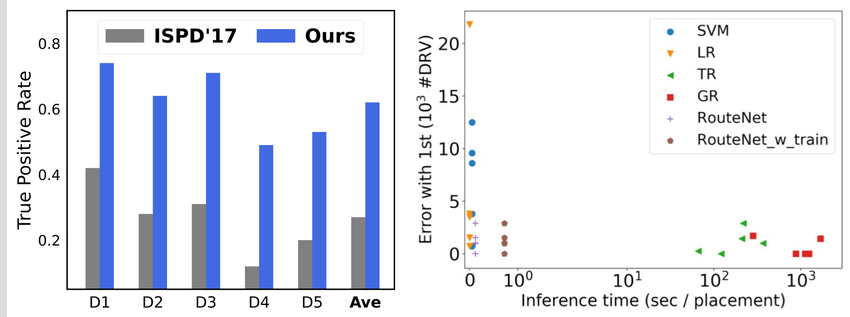
Key Method

- Define **novel** features capturing macro and long interconnect impact
- 3D input tensor constructed by stacking multiple 2D features



Experimental Results

- **Fast** and **high-fidelity** routability prediction at the same time



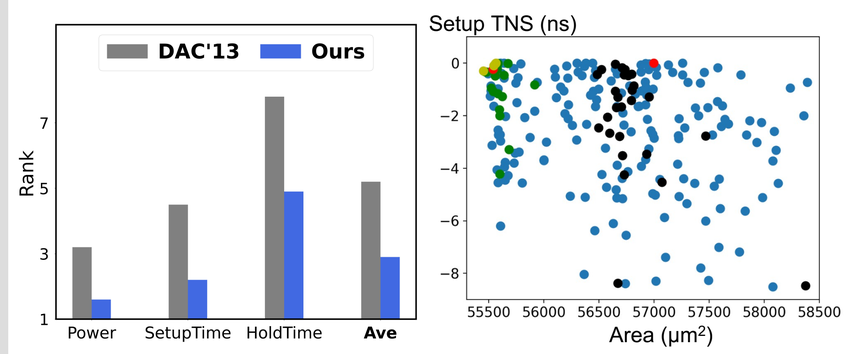
Parameter Tuning for the Design Flow [ASP-DAC'20]

Highlights (my work FIST)

- **First** flow tuning method leveraging prior data from other designs
- An approximate sampling strategy which leverages the idea of semi-supervised learning

Experimental Results

- **1.8% area improvement** on industrial design compared with best solutions hand-tuned by designers



Selected Related Publications

My Publications Presented in This Poster

- Zhiyao Xie, et al. "APOLLO: An Automated Power Modeling Framework for Runtime Power Introspection in High-volume Commercial Microprocessors." In **MICRO**, 2021. **(Best Paper Award)**
- Zhiyao Xie, et al. "Pre-placement Net Length and Timing Estimation by Customized Graph Neural Network." In **TCAD**, Under-review.
- Zhiyao Xie, et al. "Net²: A Graph Attention Network Method Customized for Pre-placement Net Length Estimation." In **ASP-DAC**, 2021.
- Zhiyao Xie, et al. "PowerNet: Transferable dynamic IR drop estimation via maximum convolutional neural network." In **ASP-DAC**, 2020.
- Zhiyao Xie, et al. "Fast IR Drop Estimation with Machine Learning." In **ICCAD**, 2020.
- Zhiyao Xie, et al. "FIST: A Feature Importance Sampling and Tree-based Method for Automatic Design Flow Parameter Tuning." In **ASP-DAC**, 2020.
- Zhiyao Xie, et al. "RouteNet: Routability Prediction for Mixed-size Designs using Convolutional Neural Network." In **ICCAD**, 2018.

My Other Publications on This Topic

- Chen-Chia Chang, Jingyu Pan, Tunhou Zhang, Zhiyao Xie, et al. "Automatic Routability Predictor Development using Neural Architecture Search." In **ICCAD**, 2021.
- Rongjian Liang, Zhiyao Xie, et al. "Routing-free Crosstalk Prediction." In **ICCAD**, 2020.
- Yu-Hung Huang, Zhiyao Xie, et al. "Routability-driven Macro Placement with Embedded CNN-based Prediction Model." In **DATE**, 2019.

Contact me for any further discussion: zhiyao.xie@duke.edu

Zhiyao Xie – ECE Department – Duke University **Email: zhiyao.xie@duke.edu**