# Reinforcement Learning-based Auto-router considering Signal Integrity

Minsu Kim, Hyunwook Park, Seongguk Kim, Keeyoung Son, Subin Kim, Kyunjune Son, Seonguk Choi, Gapyeol Park and Joungho Kim
*School of Electrical Engineering*
*Korea Advanced Institute of Science and Technology*
Daejeon, South Korea
E-mail : min-su@kaist.ac.kr

***Abstract*—In this paper, we propose artificial intelligent (AI)-router, a reinforcement learning (RL)-based auto-router considering signal integrity (SI), for the first time. Our algorithm has two main stages. At first, we design the transformer-based novel neural architecture considering the keep-out region, crosstalk region, and the number of vias for SI optimization. Then, the designed neural network is optimized by the policy gradient, one of the RL algorithms. Compared with the conventional maze routers, the A* algorithm, and the lee algorithm, it is verified that our AI-router outperforms the algorithms in terms of wire-length and crosstalk in a specific test case. Furthermore, it is shown that AI-router successfully performs multi-layer routing which is not feasible with conventional maze routers.**

***Index Terms*—AI-router, reinforcement learning, signal integrity, transformer**

## I. Introduction

Recent technology trends require terabytes per second bandwidth. To meet these demands, not only the data rate but also channel density has been significantly increased. Moreover, high integration and miniaturization of modern electronic devices critically increase design complexity in channel routing. Therefore, automatic routing considering wire-length, crosstalk, via, and layer selection became a more challenging problem.

Previous researches on auto-routing mainly consist of two stages that global routing strategies and detailed routing [1]. Global routing strategies provide a feasible routing plan including routing order. After that, each pin-to-pin routing is completed by avoiding obstacles using a maze router. However, existing maze routers have two major limitations. Firstly, existing maze routers are not smart enough to perform detailed routing considering SI [2]-[3]. Because the maze-routers cannot consider crosstalk and vias, several hand-crafted heuristics have to be designed to control the maze router for ensuring SI which leads to increases entire algorithm's complexity. Secondly, existing maze-routers have the disadvantage of modifying the entire algorithm for solving the problem which contains additional constraints and optimization variables.

On the other hands, it has been proven several times that learning-based methods can be used universally in a variety of situations without major modifications to the algorithm. In particular, recent studies have suggested that reinforcement learning shows effective results in combinatorial optimization problems including routing and hardware design [4]-[5].

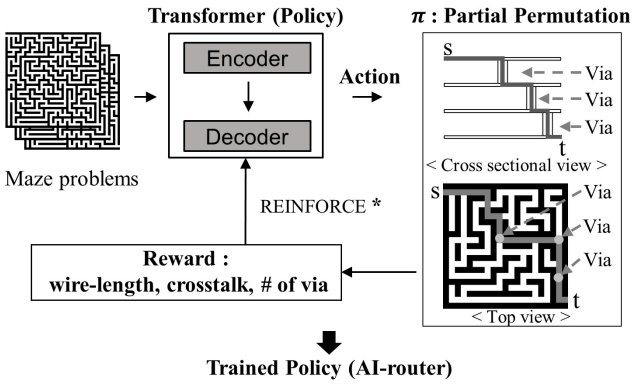


Fig. 1. Overview of training AI-router by RL. Randomly generated maze problems are training data for transformer-based encoder-decoder, which is a policy network. Then the policy network decides action which makes the partial permutation of input coordinate. The value of action is evaluated by the reward function which contains wire-length, crosstalk, and number of vias.

In this paper, we propose RL-based auto-router considering SI to overcome the limitation of the previous maze routers in terms of SI and versatility. We construct the encoder-decoder model based on the transformer, a state-of-the-art neural network in natural language processing and learning-based combinatorial optimization [5]-[6]. The policy gradient is used for training the encoder-decoder model [7]. For verification, the routing performance of the lee algorithm, the A* algorithm, and the proposed AI-router is compared in terms of wire-length and crosstalk [2]-[3]. Furthermore, we demonstrate that the proposed AI-router performs multi-layer routing which is infeasible to conventional maze routers.

## II. Proposal of Reinforcement Learning-Based AI-router

As shown in Fig. 1, a randomly generated maze problem is used as training data for the transformer-based encoder-decoder model. Through via, it is possible to move to another layer, then another maze problem of the changed layer is provided. RL is used for training policy networks which determines an optimal routing path. The trained policy network eventually becomes AI-router. Table I contains variables for the equations described in section II.A and II.B.

978-1-7281-6161-7/20/$31.00 ©2020 IEEE

TABLE I
EXPLANATION OF VARIABLES FOR EQUATIONS.

| Variables | Representation | Description |
|---|---|---|
| $X$ | $\{\mathbf{x} = (x, y, \gamma z, m_o, p_c, p_l)\}$ | Input coordinates |
| $m_0$ | $m_0 = \begin{cases} 1 & (x,y,z) \in O \\ 0 & otherwise \end{cases}$ | Obstacle mark |
| $p_c$ | $p_c = \begin{cases} 2 & (x,y,z) \in C \\ 0 & otherwise \end{cases}$ | Penalty of Xtalk |
| $p_l$ | $p_l = \begin{cases} 0.5 & (x,y,z) \in L \\ 0 & otherwise \end{cases}$ | Penalty of loosely Xtalk |
| $\gamma$ | $\gamma = 30$ | Layer changing penalty |
| $\pi$ | $\{\pi_1, \pi_2, ..., \pi_n\}$ , $x_{\pi_i} \in X$ | Output Permutation |
| $O$ | $\{(x_i, y_i, z_i)\}$ | Obstacle region |
| $C$ | $\{(x_i, y_i, z_i)\}$ | Xtalk region |
| $L$ | $\{(x_i, y_i, z_i)\}$ | Loosely Xtalk region |

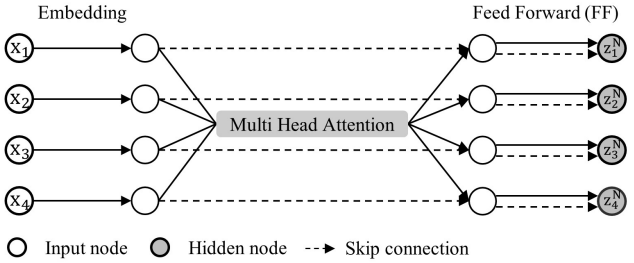


Fig. 2. Conceptual view of encoder architecture. Input nodes **x** are embedded to high dimensional vectors $z$ by multi-head attention, feed forward, and skip connection.

### A. Neural Architecture of AI-router

As illustrated in Fig. 2, the encoder architecture mainly consists of multi-head attention (MHA), feed forward (FF), and skip connection (Skip) which embed input node to high dimensional hidden nodes $z^N$. A detailed equation of the Skip, MHA, and FF is explained in [5]-[6].

First, the input node is embedded through the linear projection as the following equation:

$$z_i^0 = W\mathbf{x}_i + b \tag{1}$$

$W$ and $b$ are learnable parameters. Then, the hidden nodes $z^N$ are made from the embedded node with MHA as the following equation:

$$z_j^{i+1} = Skip_j(FF(Skip_j(MHA_j(z_1^i; z_2^i; :::; z_n^i)))) \tag{2}$$

Then the hidden nodes are propagated forward the decoder as illustrated in Fig.3. The main purpose of the decoder is to output the stochastic policy. To be specific, *context vector* $c$ creates a policy for the next selection, referring to the previous node, the average node of all, the average node of obstacles(keep-out region and pre-routed region), the average node of coupled crosstalk region, the average node of loosely-coupled crosstalk region and the destination(target pin) node:

$$c_k = F_1([z_{k-1}^N; z_n^N]) + F_2([z_{mean}^N; z_o^N; z_c^N; z_l^N]) \tag{3}$$

$F_1$ and $F_2$ is fully-connected layer with learnable parameter $_1$, $_2$. As shown in (5), the context vector passes MHA and becomes query $c'$.

$$c'_k = MHA(z_1^N; :::; z_n^N; c_k) \tag{4}$$

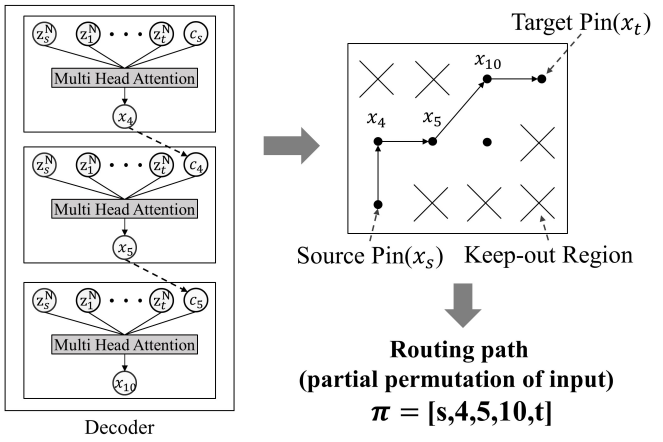


Fig. 3. Conceptual view of proposed decoder processing. The decoder uses the context vector $c$ and MHA to determine the next node. The determined node affects the selection of a next node by changing the context vector $c$. The decoder eventually creates a routing path of $\pi$.

Then the partial policy of selecting the next action can be expressed as follows:

$$u_j = \begin{cases} 10\tanh(\frac{(Qc')^t(Kz_j)}{4}) & nodes\ 2\ Action \\ 1 & otherwise \end{cases} \tag{5}$$

$$P(\ _i = j|X;\ _{i-1};\ _n) = \frac{e^{u_j}}{\sum_{k=1}^{n} e^{u_k}} \tag{6}$$

Q and K in (5) are learnable parameters which make query and key, respectively. A detail of the query and key is introduced in [5]-[6]. The $Action$ in (5) contains left, right, up, down, 45-degree routing and layer changing through via. Also, obstacles are not possible to be selected by the $Action$.

Therefore the policy of routing permutation is

$$P(\ |X) = \prod_{k=2}^{n-1} P(\ _k|X;\ _{k-1};\ _n) \tag{7}$$

### B. Training the AI-router with Reinforcement Learning

The policy gradient is used to train the AI-router architecture [7]. When we get permutation of nodes , we calculate the cost of AI-router as follows:

$$L(\ ) = \sum_{k=1}^{n-1} \left(g(\mathbf{x}_{\ k+1} \quad \mathbf{x}_{\ k}) + p_c(\mathbf{x}_{\ k}) + p_{lc}(\mathbf{x}_{\ k})\right) \tag{8}$$

$$g(\mathbf{x}) = \sqrt{x^2 + y^2 + \ ^2z^2} \tag{9}$$

To estimate the crosstalk in the cost function $L(\ )$, we set the crosstalk region and loosely crosstalk region at 1$w$ and 2$w$ distances based on the pre-routed channel, respectively. The $w$ is the width of the routing channel. In the end, the total crosstalk is estimated by counting the number of blocks that are invading crosstalk regions on a grid. A block in the grid is a path which is made by node selection from the $\mathbf{x}_{\ k}$ to $\mathbf{x}_{\ k+1}$ in (8).

Also, we assign as a penalty for layer change. When changing a layer, it greatly interferes with the routing plan of other layers. Also, the via transition effect adversely affects SI. Therefore, layer change through via is not recommended