

Research Progress Report

Botao Zhu

January 5, 2019

1 Reading and Research Activities

1.1 Summary

[1] proposed a deep learning based routing table construction method for a GPU-accelerated SDR. They adopted supervised Deep Belief Architectures (DBA) to compute the subsequent nodes with the traffic patterns of the edge routers as the input. And then, they presented unique characterizations of inputs and outputs based on the traffic patterns at the edge routers. They demonstrated how the trained DBA can predict the nodes, the benefits of the deep learning based routing strategy in terms of lower signaling overhead as well as fast convergence and effectiveness of proposed deep learning based solution compared to a benchmark routing method through both analysis and extensive.

1.2 Introduction

1.2.1 Deep Learning Based Routing Strategy

1.2.1.1 Input and Output Design The traffic pattern is served as the input to the deep learning structure and processed for routing path decision as the output. The traffic pattern at each router can be defined as the number of inbound packets of the router during each time interval, such as $\beta\Delta t$. Therefore, by assuming that a network comprises of N routers, we can use a matrix of β rows and N columns to represent the traffic patterns of all routers. The simulation results demonstrate that it is accurate enough to set the value of β to 1. The output layer can be designed to give the next node similar to the distributed routing strategy because of its lower complexity and higher tolerance. In the vector, only a single element has the value of 1, the order of which represents the next node. So, we can use two N dimensional vectors to represent the input and output.

1.2.1.2 Deep Learning Structure Design The authors choose the DBA as the deep learning structures as shown in Figure 1.

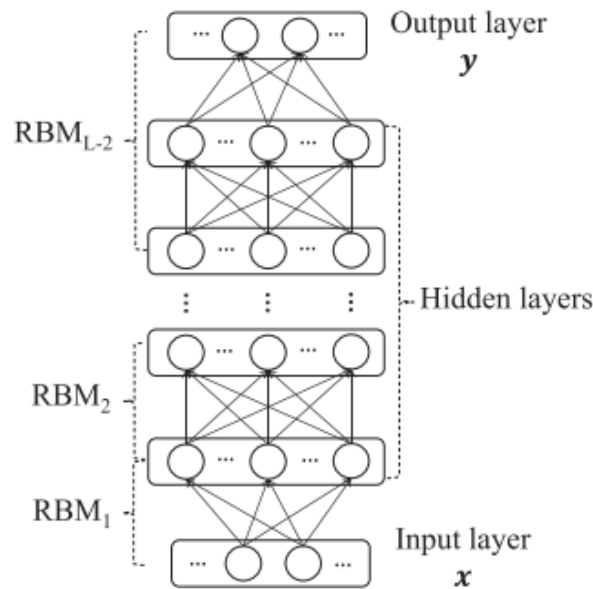


Figure 1: Considered L-layer DBA

DBA can be also seen as a stack of $(L - 2)$ Restricted Boltzmann Machines (RBM) and a logistic regression layer as the top layer. The structure of RBM consists of two layers, the visible layer, \mathbf{v} ,

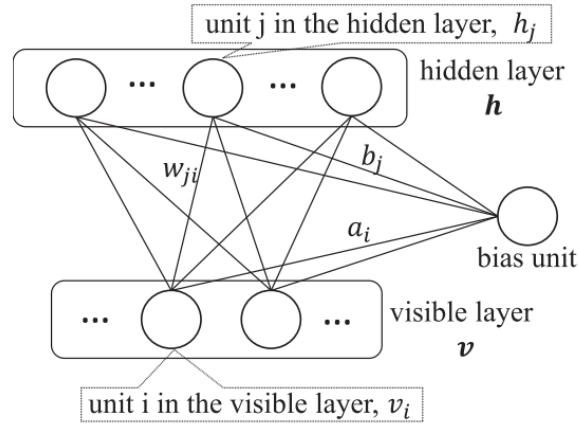


Figure 2: RBM

and the hidden layer, \mathbf{h} , shown in Figure 2. While training an RBM, sets of unlabeled data are given to the visible layer, and the value of the weights and biases are repeatedly adjusted until the hidden layer can reconstruct the visible layer. To mathematically model the training process, they use a log-likelihood function of the visible layer given as follows.

$$l(\theta, a) = \sum_{t=1}^m \log p(v^t) \quad (1)$$

where θ denotes the vector consisting of all the values of the weights and biases of the hidden layer. Since RBM is a particular form of log-linear Markov Random Field, the energy function, (v, h) , and the joint probability function, $p(v, h)$, are defined as follows:

$$E(v, h) = - \sum_i a_i v_i - \sum_j b_j h_j - \sum_i \sum_j h_j w_{ji} v_i \quad (2)$$

$$p(v, h) = \frac{e^{-E(v, h)}}{Z} \quad (3)$$

$$Z = \sum_v \sum_h e^{-E(v, h)} \quad (4)$$

where v_i and h_j are the unit i in the visible layer and the unit j in the hidden layer respectively. Z represents the normalizing constant partition function. Also, the relationship between $p(v)$ and $p(v, h)$ can be expressed as follows:

$$p(v) = \sum_h p(v, h) \quad (5)$$

However, the complexity of the calculation is extremely high. To solve this problem, they use the Gibbs Sampling method to sample the values of v and h to approximate the real values since the conditional distribution probability of one layer can be calculated. As each unit is independent on the other units in the same layer, when one layers is fixed, the conditional distribution probability of the other layer can be calculated as follows:

$$p(v|h; \theta, a) = \prod_i p(v_i|h; \theta, a) \quad (6)$$

$$p(h|v; \theta, a) = \prod_j p(h_j|v; \theta, a) \quad (7)$$

Since input units are continuous, they use Gaussian probability distribution to model the traffic patterns. Equations(2) and (6) should be revised as follows:

$$E(v, h) = -\sum_i \frac{(v_i - a_i)^2}{2\delta_i^2} - \sum_j b_j h_j - \sum_i \sum_j \frac{v_i}{\delta_i} h_j w_{ji} \quad (8)$$

$$p(v_i|h; \theta, a) = N\left(a_i + \delta_i \sum_j h_j w_{ji}, \delta_i^2\right) \quad (9)$$

where δ_i is the value of the variance for the unit v_i . However, in Figure 3, we can see that the visible layer of RBM_{L-2} consists of not only RBM_{L-3} 's hidden layer but also the output layer of the DBA, y . Its hidden layer is the top hidden layer of the DBA and energy function is expressed as follows.

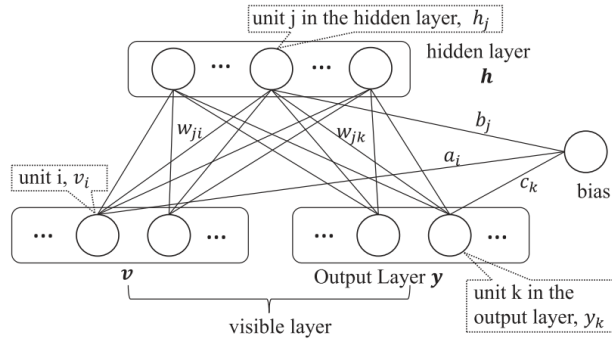


Figure 3: Last RBM

$$E(v, h, y) = -\sum_i a_i v_i - \sum_j b_j h_j - \sum_k c_k y_k - \sum_i \sum_j h_j w_{ji} v_i - \sum_j \sum_k h_j w_{jk} y_k \quad (10)$$

where y represents the vector in the output layer. c_k is the bias of the unit y_k . w_{jk} represents the weight of the link connecting the units h_j and y_k . The conditional distribution of the concatenated vector consisting of v and y is,

$$p(v, y|h; \theta, a) = p(v|h; \theta, a) p(y|h; \theta, a) \quad (11)$$

$$= \prod_i p(v_i|h; \theta, a) \prod_k p(y_k|h; \theta, a) \quad (12)$$

2 Objectives for the Next 2 Weeks

3 Advisor's Comments

I'm very happy.cite [1]

Bibliography

- [1] (2015) The IEEE website. [Online]. Available: <http://www.ieee.org/>