# Experiments

2018 3.25

Modify action space

- We suppose that each flow has some valid(following BGP advertisement) next-hops to choose as action. In the previous setting of RL, we consider the combination of all flows'(in traffic matrix) valid next-hops to be the action space, which will cause an exponentially growth in size.

- Hence, we simplify the definition of action space as neighboring nodes set, whose size only relates to the degree of agent.

- Specifically, when agent receives state and choose an action based on policy distribution(output layer of policy network) on action space, it masks invalid action(not following BGP advertisement) by setting its probability as 0. Then agent rescales the distribution such that the sum still equals to 1 and chooses action from new distribution.

Experiment results

- On 30 nodes, 24 flows network with optimal system throughput value 1383.3
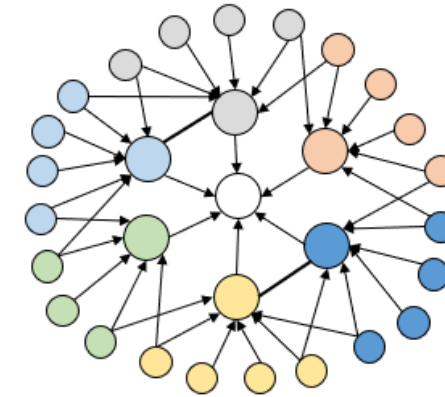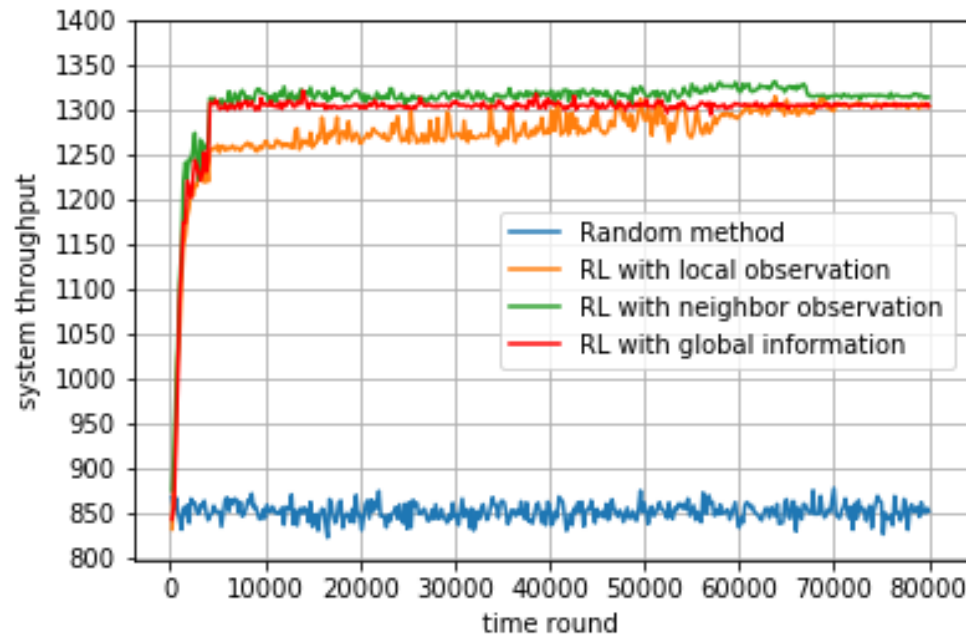
The system throughput at the stable state still has micro oscillation. E.g. between the value 1316.6 and 1300.0.

- All the three situations can reach state with system throughput above 1300. For local observation, it traps into local optimal around 1250 at 10000 time round, but still gradually improves in later time.

- Try dynamic TF pattern.
    A FIFO queue maintains 20 flows, each time epoch we choose 10 flows to be the traffic matrix and train it for a whole epoch time (1000 time rounds).
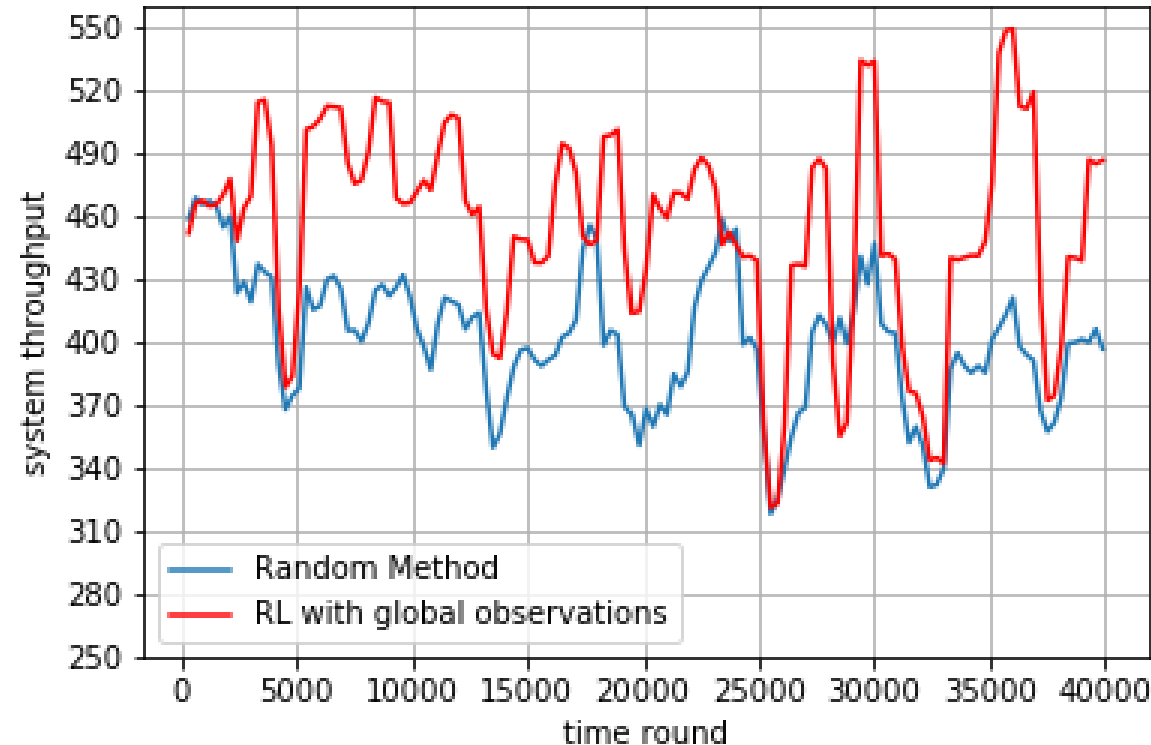
RL model can have little improvement(average 15%) than random method, but may be difficult to get into a stable stage for each incoming traffic matrix.