Packet Routing with Hierarchical Reinforcement Learning

August 7, 2022

1 Network Topology

Take the network topology shown in Fig. 1 as an example. All nodes are divided into 4 groups, named *Group A, B, C, D*, and each group includes different number of nodes, e.g., *Group A* includes 3 nodes and node a_1 is the first node belonging to *Group A*. the nodes in the network are grouped (clustered) based on specific criteria such as the latency between the nodes or geographical positions. Furthermore, readers are reminded that a requirement is put on the network topology. that is, if node i does not belong to a group, there is at most one edge between node i and the nodes belonging to that group.

2 Hierarchical Routing

Our hierarchical routing includes two parts: routing among groups and routing within groups. The routing strategy for each part is decided by two Q-tables, called *outer Q-table* for routing among groups and *inner Q-table* for routing within groups. Each group maintain a *outer Q-table*, and the *outer Q-table* is shared and updated by all nodes belonging to the corresponding group. However, each node maintain its own *inner Q-table*. For example, Fig. 2 and Fig. 3 show the *outer Q-table* of *Group A* and the *inner Q-table* of node a_1 respectively. In *outer Q-table*, the entry indicates the expected time-to-arrival that it takes to route a packet from current group to destination group by way of a particular next group hop. For example, in Fig. 2, the expected time-to-arrival from *group A* to *group D* by way of b_3 is 11. While *inner Q-table* is more complex and the entry is responsible for two different routing cases: one is for the packet that is bound for next group hop by way of a particular next hop, shown in the upper half in Fig. 3. For instance, the expected time-to-arrival for the packet in node a_1 bound for b_3 by way of a_2 is 4. The other is for the packet that arrives at the destination group, shown in the lower half in Fig. 3. "\" means that once the packet arrives at the destination group, it will not jump out.

Next, we firstly introduce the routing strategy by an example where a packet need to be transmitted from source a_1 to destination d_3. Assume that the optimal transmission route

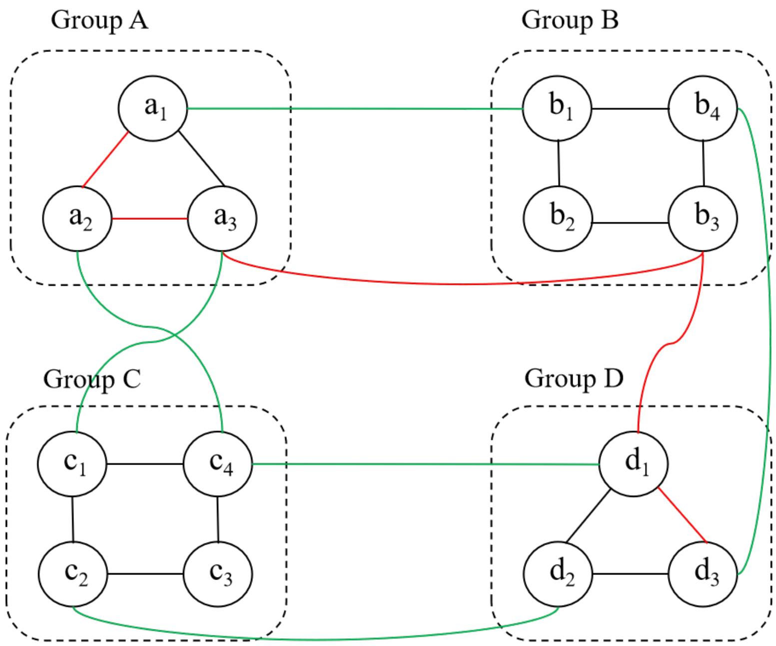

Figure 1: Network topology

is $a_1 \rightarrow a_2 \rightarrow a_3 \rightarrow b_3 \rightarrow d_1 \rightarrow d_3$, indicated in red line in Fig. 1. When a packet is injected to node a_1, a_1 will firstly verify whether the packet arrives at a new group. If yes, a_1 will confirm whether current group is the destination group of the packet. If no, a_1 will search its own *inner Q-table* to find the optimal next hop. Since the destination group is *Group D*, a_1 will search the *outer Q-table* of *Group A* and find that the optimal next group hop is b_3. Then a_1 will search its own *inner Q-table* to find the optimal next hop to b_3. Since $(b_3, *)$ entry is minimal when $*$ is a_2, then a_1 will allocate the packet to a_2. After the packet arrives at a_2, a_2 will also confirm whether it arrive at a new group. Since the packet is still in *Group A*, a_2 will search its own *inner Q-table* to find the optimal next hop to b_3. Once b_3 receives this packet, it will repeat the process experienced in node a_1.

Based on the above example, we summarize the routing strategy. Fig. 4 presents the corresponding flow chart for hierarchical routing. Please note that the *outer Q-table* will be used only when a packet arrives at a new group and that group is not the destination group of that packet.

3 Update of *outer Q-table* and *inner Q-table*

In this part, we will introduce the update strategy for *outer Q-table* and *inner Q-table*.

A		Next group hop			
		b_1	b_3	c_1	c_4
Dest. Group	B	...	...	...	...
	C	...	...	...	...
	D	21	11	17	24

Figure 2: Outer Q-table of Group A

a_1		Next hop		
		a_2	a_3	b_1
Next group hop	b_1	...	...	...
	b_3	4	6	8
	c_1	...	...	...
	c_4	...	...	...
If packet in dest. group	a_2	0	...	\
	a_3	...	0	\

Figure 3: Inner Q-table of node a_1

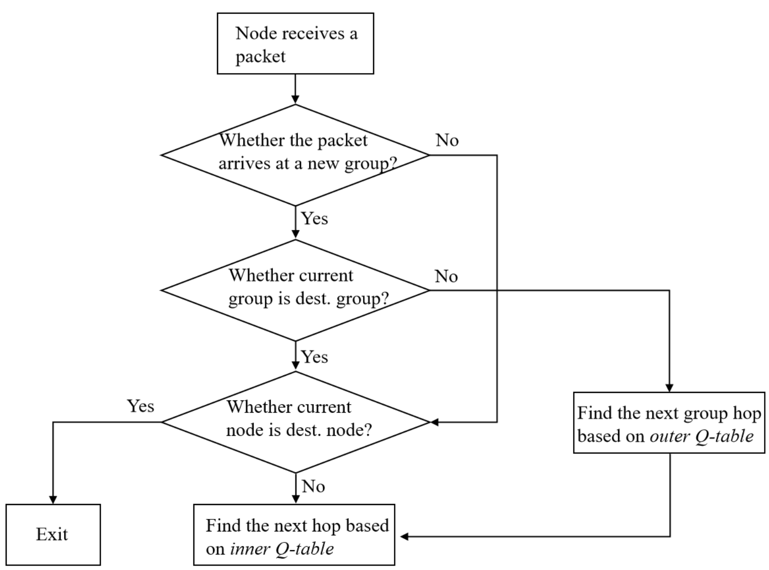

Figure 4: Flow chart for hierarchical routing

3.1 Update of *outer Q-table*

Firstly, let $Q_X(Z, y)$ be the time that an *outer Q-table* of *Group* X estimates it takes to deliver a packet P bound for destination group Z by way of next group hop y. Upon sending P to y, *Group* X immediately gets back y's estimate for the time remaining in the trip. For instance, in Fig. 2, $Q_A(D, b_3) = 11$.

Based on the above described routing strategy, we firstly indicate that the *outer Q-table* will be updated when a packet leaves a group and enters a new group. Since the entry in *outer Q-table* indicates the expected time-to-arrival that it takes to route a packet from current group to destination group by way of a particular next group hop, we need to know the total travelling time from the time slot when a packet arrives at a new group to the time slot when the packet arrives at the next new group. We denote this duration of time as $R_{X\to Y}$, where X, Y are the group ID, e.g., $X, Y \in \{A, B, C, D\}$. For example, The duration of time from the time slot when a packet arrives at a_1 to the time slot when the packet reaches b_3 is $R_{A\to B}$. Thus, the update of *outer Q-table* is:

$$Q_X(Z, y) = R_{X\to Y} + argmin_s(Q_Y(Z, s)), y \in Y$$

3.2 Update of *inner Q-table*

Similarly, we define $Q_x(z, y)$ as the time that a node x estimates it takes to deliver a packet P bound for destination node z by way of x's neighbor node y, including any time that P would have to spend in node x's queue. Upon sending P to y, x immediately gets back y's estimate for the time remaining in the same group. For example, in Fig. 3, $Q_{a_1}(b_3, a_2) = 4$. What's more, let $r_{x\to y}$ indicate the travelling time spent from x to y, i.e., the sum of queuing time in x and transmission time between node x and y. Then, the update of *inner Q-table* is:

$$Q_x(z, y) = r_{x\to y} + argmin_s(Q_y(z, s))$$