Consideration on Applying Q-Learning to Backpressure Routing Algorithm to Improve Delay Performance

JUNTAO GAO†1,a)　MINORU ITO†1,b)

Back-pressure routing algorithm becomes increasingly popular for queueing networks, such as wireless ad hoc networks and MANETs. However, it is well known that backpressure routing algorithm has poor delay performance under light and moderate traffic loads. Available works propose to exploit shortest path and global queue length information to direct packets to shorter routes to their destinations to reduce packet delay. However, shortest path based backpressure routing inclines to route packets to shortest paths regardless of traffic congestion, thus causing long packet delay as traffic load increases. Global queue length based backpressure routing algorithm requires perfect knowledge of global queue length information which is hard to collect in real queueing networks. In this paper, we propose a Q-learning based backpressure routing (QL-BP) algorithm, which estimates route congestion based on only local queue length information. Our algorithm cannot only dynamically direct packets to routes of less congestion and effectively reduce packet delay, but also retain all appealing features of backpressure routing: throughput-optimality, distributed implementation and low computational complexity. As verified by simulations, our QL-BP algorithm reduces average packet delay by more than 71% when compared to traditional BP algorithm under light and moderate traffic loads. By enhancing QL-BP algorithm with shortest path information, our QL-BP algorithm outperforms all other variants of backpressure routing algorithms.

1. Introduction

Since the seminal work [1], back-pressure routing algorithm becomes increasingly popular for queueing networks, like wireless ad hoc networks and MANETs, due to its appealing features: throughput-optimality, distributed implementation and low computational complexity. Back-pressure routing algorithm explores all possible routes to balance traffic across the whole queueing network to accommodate as much traffic as possible, thus achieving throughput optimality. To achieve traffic balance, each node under backpressure routing algorithm communicates with its neighboring nodes one hop away for queue length information and transmits packets of queues with the largest queue differential to its neighboring nodes [2].

However, the myopic feature (considering only queue length of neighboring nodes) of backpressure routing algorithm makes nodes blind to queue length information of farther nodes (i.e., nodes multi-hop away) and destination locations. This leads to directing packets to unnecessarily long routes or event route loops under light and moderate traffic loads as shown in Fig. 1, which finally results in long packet delay.

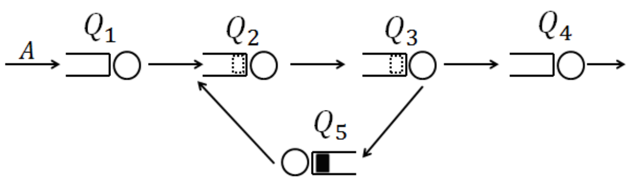

Fig. 1 A packet being directed to a route loop

Available works propose to exploit shortest path and global queue length information to direct packets to shorter routes to their destinations to reduce packet delay [2,3,4,5]. Although shortest path based backpressure routing algorithm performs well in delay performance under light traffic loads, it still results in long packet delay for heavier traffic loads since it inclines to route packet to shortest paths, causing traffic congestion. The state of the art algorithm for reducing packet delay of backpressure routing is called BPmin proposed in [3], which directs packets to the least congested routes, i.e., routes with least number of packets along that route. However, under BPmin algorithm, every node needs to know queue length information of all other nodes in the network to evaluate the degree of route congestion. Such global queue length information is hard to collect in realistic queueing networks.

To address this problem, we propose in this paper a Q-learning based backpressure routing algorithm, where each node evaluates route congestion degree based on only queue length information of neighboring nodes. Our algorithm cannot only dynamically route packets to paths of less congestion thus effectively reducing packet delay, but also retain all appealing features of backpressure routing: throughput-optimality, distributed implementation and low computational complexity.

2. Q-Learning Based Backpressure (QL-BP) Routing Algorithm

Consider a multi-hop queueing network represented by a directed graph $\mathcal{G} = (\mathcal{N}, \mathcal{L})$, where $\mathcal{N}$ is the set of N nodes and $\mathcal{L}$ is the set of L directed links. Time is slotted. Let $U_i^{(c)}(t)$ be the number of packets destined for node c and queued up at node i at slot t. Packets destined for node c are referred to as commodity c.

To evaluate the degree of route congestion, each node i maintains a table $\boldsymbol{Q}_i^{(c)}$ for each commodity c, where each entry $Q_{ij}^{(c)}$ records the estimate of the congestion degree of routes starting from link (i,j). Under Q-learning based backpressure

†1 Graduate School of Information Science, Nara Institute of Science and Technology, Japan
a) jtgao@is.naist.jp
b) ito@is.naist.jp

©2017 Information Processing Society of Japan

routing algorithm, each node *i* executes the following every time slot .

Collect information: At every time slot *t*, node *i* exchanges information of queue length $U_i^{(c)}(t)$ and tables of route congestion estimate $\boldsymbol{Q}_i^{(c)}$ with its neighboring nodes.

Q-learning based congestion estimate: Node *i* updates its congestion estimate $Q_{ij}^{(c)}$ for all commodities c and links (i,j) based on Q-learning as follows

$$Q_{ij}^{(c)} \longleftarrow (1-\alpha)Q_{ij}^{(c)} + \alpha\left[U_j^{(c)}(t) + \gamma \min_{k\in\text{neighbor of } j} Q_{jk}^{(c)}\right] \tag{1}$$

where α and γ are Q-learning parameters, $0 < \alpha, \gamma < 1$.

Then, node *i* calculates the congestion degree $V_i^{(c)}$ of routes starting from node *i* for commodity *c*.

$$V_i^{(c)} = \min_j Q_{ij}^{(c)} \tag{2}$$

Backpressure routing:

Node i finds the optimal commodity $c_{ij}^*(t)$ for every link (i,j) such that

$$c_{ij}^*(t) = argmax_{c\in\mathcal{N}}\left[\left(U_i^{(c)}(t) + V_i^{(c)}\right) - \left(U_j^{(c)}(t) + V_j^{(c)}\right)\right] \tag{3}$$

Then, node *i* calculates pressure gradients $W_{ij}^*(t)$ for every link (i,j)

$$W_{ij}^*(t) = \max\left[\left(U_i^{(c_{ij}^*(t))}(t) + V_i^{(c_{ij}^*(t))}\right) - \left(U_j^{(c_{ij}^*(t))}(t) + V_j^{(c_{ij}^*(t))}\right), 0\right] \tag{4}$$

Node *i* makes resource allocation decision $I_i^{\mathrm{QL-BP}}(t)$ such that

$$I_i^{QL-BP}(t) = \mathrm{argmax}_{I_{ij}(t),(i,j)} \sum_{(i,j)} \mu_{ij}\left(S_{ij}(t), I_{ij}(t)\right) W_{ij}^*(t) \tag{5}$$

where $S_{ij}(t)$ is the channel state of link (i,j) at slot t and $I_{ij}(t)$ is the resource allocation decision for link (i,j) at slot t, e.g., link activation, coding, modulation, μ_{ij} is transmission data rates of link (i,j), depending on $S_{ij}(t)$ and $I_{ij}(t)$.

Finally, node *i* transmits packets with following data rates for commodity c and link (i,j)

$$\mu_{ij}^{(c)QL-BP}(t) = \mu_{ij}\left(S_{ij}(t), I_i^{QL-BP}(t)\right), \quad \text{if } c = c_{ij}^*(t) \text{ and } W_{ij}^*(t) > 0$$

$$\mu_{ij}^{(c)QL-BP}(t) = 0, \quad \text{otherwise} \tag{6}$$

3. Simulation

We evaluate the delay performance of our algorithm by simulations and compare it to other variants of backpressure routing algorithms.

3.1 Simulation Setup

We consider the network topology as shown in Fig. 2, which consists of 64 nodes, indexed by a pair of coordinates. All links are bidirectional and the maximum data transmission rates for all links are 1 packet/slot. We assume all links can transmit packets simultaneously without interfering with each other, such as wireline network or wireless network with orthogonal channels. We consider traffic flows with the following source-destination pairs: ((1,3),(2,5)), ((2,3),(2,7)), ((2,2),(1,6)), ((3,4),(2,7)), ((1,1),(1,7)), ((4,3),(5,4)), ((4,6),(6,6)), and ((5,3),(5,6)). All source nodes generate packets according to Poisson distribution with rate λ packets/slot. We implemented by Python our Q-learning based backpressure routing algorithm (QL-BP), traditional backpressure routing algorithm (BP) [2], shortest path based backpressure routing algorithm (SP-BP) [2], state-of-the-art BPmin algorithm [3]. We also implemented an enhanced version of our QL-BP by considering shortest path information (called QLSP-BP), which is the same with QL-BP except that it finds the optimal commodity $c_{ij}^*(t)$ as follows

$$c_{ij}^*(t) = argmax_{c\in\mathcal{N}}\left[\left(U_i^{(c)}(t) + V_i^{(c)} + P_i^{(c)}\right) - \left(U_j^{(c)}(t) + V_j^{(c)} + P_j^{(c)}\right)\right]$$

where $P_i^{(c)}$ is the length of the shortest path from node *i* to node *c*.

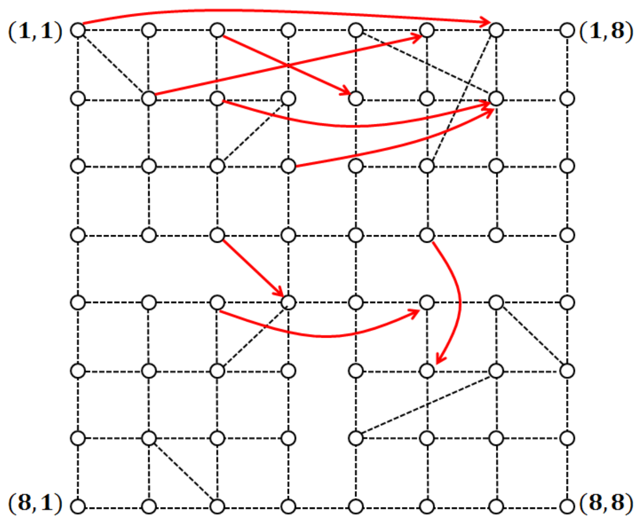

Fig. 2. Network topology for simulation

We run simulations for 10^5 slots for each packet arrival rate and calculate the average delay of packets received by destinations.

3.2 Simulation Results

From Fig. 3 we can observe that our QL-BP algorithm reduces packet delay by 71% when compared to traditional BP algorithm under light traffic load with $\lambda = 0.1$ and by 87% under moderate traffic load with $\lambda = 0.6$, indicating that QL-BP effectively learns congestion degree of different routes and adaptively direct packets to shorter routes. However, QL-BP algorithm results in slightly higher packet delay than state-of-the-art BPmin algorithm. This is because nodes of

BPmin algorithm know perfect global queue length information and thus can accurately calculate congestion degree of different routes and direct packets to the least congested routes. While nodes of QL-BP algorithm only know queue length information of neighboring nodes, thus can only estimate congestion degree of different routes, which may lead to directing packets to suboptimal routes. However, BPmin is not realistic since global queue length information is hard to collect by nodes in real world. QL-BP slightly sacrifices packet delay for distributed algorithm implementation and thus can be easily deployed in real queueing networks.

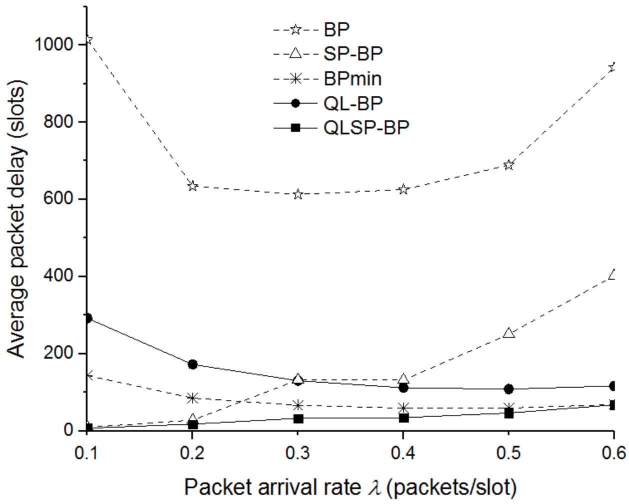

Fig. 3. Average packet delay under light and moderate traffic loads

Our QL-BP algorithm can be further improved by considering shortest path information. From Fig. 3, we see that QLSP-BP algorithm outperforms all variants of backpressure routing algorithms including state-of-the-art BPmin algorithm.

4. Conclusion

In this paper, we proposed a Q-learning based backpressure routing (QL-BP) algorithm to improve delay performance of traditional backpressure routing algorithm under both light and moderate traffic loads. As verified by simulations, QL-BP algorithm effectively reduces packet delay and outperforms state-of-the-art BPmin algorithm. Our QL-BP algorithm also retains appealing features: throughput-optimality, distributed implementation and low computational complexity. QL-BP algorithm can be further improved by considering shortest path information.

Acknowledgments

This work is supported by JSPS KAKENHI Grant Numbers JP15K15981

Reference

[1] L. Tassiulas and A. Ephremides, “Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks,” *IEEE Transactions on Automatic Control*, vol. 37, no. 12, 1992.

[2] M. J. Neely, E. Modiano, and C. E. Rohrs, “Dynamic power allocation and routing for time-varying wireless networks,” IEEE Journal on Selected Areas in Communications, vol. 23, no. 1, pp. 89–103, January 2005.

[3] Y. Cui, E. M. Yeh, and R. Liu, “Enhancing the delay performance of dynamic backpressure algorithms,” IEEE/ACM Transactions on Networking, vol. 24, no. 2, pp. 954–967, April 2016.

[4] L. Ying, S. Shakkottai, A. Reddy, and S. Liu, “On combining shortest path and back-pressure routing over multihop wireless networks,” IEEE/ACM Transactions on Networking, vol. 19, no. 3, pp. 841–854, June 2011.

[5] Z. Jiao, B. Zhang, W. Gong, and H. Mouftah, “A virtual queue based back-pressure scheduling algorithm for wireless sensor networks,” EURASIP Journal on Wireless Communications and Networking, vol. 2015, no. 1, pp. 1–9, 2015.

©2017 Information Processing Society of Japan