Survey on Reinforcement Learning based Efficient Routing in SDN

Hong-Nam Quach
Chonnam National University
Gwangju, South Korea
quachhongnam1995@gmail.com

Sungwoong Yeom
Chonnam National University
Gwangju, South Korea
yeomsw0421@gmail.com

Kyungbaek Kim
Chonnam National University
Gwangju, South Korea
kyungbaekkim@jun.ac.kr

ABSTRACT

Traffic on the Internet is growing exponentially around the globe, and SDN (Software Defined Network) becomes very attractive to network research communities. The trend switching from a hardware-centered network to a software-centric network allows that networks become more flexible and efficient. SDN is getting more and more attention, especially with the recent advent of cloud and 5G technology, and many researchers are making various attempts to improve SDN performance. Also, the application of artificial intelligence technology, which applied to multiple fields, has attracted many researchers, and a great deal of research is underway. Specifically, in many fields, including the network industry, the reinforcement learning technique has gained attention. The SDN field also reflected the recent increase in research that relates reinforcement learning. In this paper, we would briefly introduce SDN and reinforcement learning technology by identifying and analyzing recent studies which use reinforcement learning for efficient routing in SDN. Implementing RL techniques in SDN routing shows that enforcing routing to dynamic networks becomes more efficient, and it can provide excellent levels of QoS while optimizing resource utilization.

CCS CONCEPTS

• **Networks → Routing protocols**.

KEYWORDS

SDN, Artificial Intelligence, Reinforcement Learning, Survey

ACM Reference Format:

Hong-Nam Quach, Sungwoong Yeom, and Kyungbaek Kim. 2020. Survey on Reinforcement Learning based Efficient Routing in SDN. In *The 9th International Conference on Smart Media and Applications (SMA 2020), September 17–19, 2020, Jeju, Republic of Korea.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3426020.3426072

1 INTRODUCTION

Nowadays, data traffic in the world is increasing significantly, along with the increased development of smart devices, network technologies. Accordingly, it is necessary to change network management technology. By switching from a hardware-oriented network to a software-based network system, it is possible to manage and control the system logically. This change has made it possible to manage distributed systems centrally. As a result, SDN is becoming an inevitable technology in distributed systems such as IoT and large data-based large networks. Thanks to the automated orchestration function of SDN, it is possible to execute more flexible and high - performance communication than the existing hardware-centric network. In other words, automatic system, storage, and server orchestration are essential for running a data center in the SDN - based cloud, and it can be achieved flexibly by having it in SDN, a software-based network, according to the situation. SDN is thus attracting attention as a promising networking technology for the next decade. Additionally, in recent years, artificial intelligence (AI) technology has been rising rapidly. AI fusion and complex have emerged in many fields, especially after the 4th Industrial Revolution. From AI assistants that are easily accessible for efficient production in everyday life to smart factories, they are rapidly grafting in diverse industries, creating significant added value. The network industry, therefore, prefers also to introduce AI technology. The SDN has become one of the areas of great interest. It can see that the combination of artificial intelligence with SDN has led various studies to improve performance. Thus, it has become an incentive to draw further interest because of the benefit that data easily obtained via SDN, which is the most critical factor in AI learning. Therefore, many researchers are actively researching improving performance by using reinforcement learning in the existing SDN, and there have been researching papers [10]. However, we can see that more and more studies have applied reinforcement learning among various learning methods to SDN. In this paper, recent studies using reinforcement learning in SDN summarized and analyzed. This paper presented the following sections. Chapter II provides background, and Chapter III is RL based efficient routing in SDN. Finally, we finish with IV.

SMA 2020, September 17–19, 2020, Jeju, Republic of Korea

ACM ISBN 978-1-4503-8925-9/20/09…$15.00
https://doi.org/10.1145/3426020.3426072

2 BACKGROUND

Brief knowledge of SDN architecture and reinforcement learning are presented in this section.

2.1 An overview of SDN

Software-Defined Networking [2] provided a dynamic, manageable, and cost-effective platform for making it an important platform for the high-bandwidth, dynamic nature of today's network application. The SDN architecture consists of three primary planes, including the data plane, the control plane, and the application plane. The Southbound-API (SB-API) provides an interface for data- and control plane interaction. There are several protocols available

Table 1: Target and learning elements in routing

Ref	Objective	Algorithm	State	Action	Reward
[3]	Optimize Routing	Q-learning	Link discovery and classification, Intensive learning training	Link weight (delay and bandwidth)	Conversational services, Streaming media service, Interactive service, Data class service
[12]	Controller synchronization	Q-learning	Time slot counts	Synchronization Decision	Average of reduced path cost
[13]	Optimize Routing	DRL	Service creation location	The number of edge node routing requests	Edge node access status, Resource usage balance degree, Request data transmission delay
[9]	Optimize Routing	DDPG	Network status	Link weight matrix	Service quality (QoS)
[11]	Optimize routing	DDPG	Traffic Matrix (TM)	Link weight	Network operation and maintenance strategy
[7]	Optimize Routing	DDPG	Traffic Matrix (TM)	Link weight vector	QoS metrics and Qualified Flows

for the two planes to communicate, such as OpenFlow and Netconf. The control plane is implemented by means of SDN controllers, such as POX, ODL, ONOS, Floodlight, and RYU. The Northbound API (NB-API) is an interface between the management and control planes. The SDN controller acts as a bridge between the management and data planes taking advantage of the Representative State Transfer API (REST). Likewise, data plane information, such as the flows obtained through the REST API. Moreover, the South Bound and North Bound APIs are provided as communication channels between the SDN planes. SDN's Programming and coordination capabilities are deployed through the application plane at the top floor of the control plane. Different applications are carried out to perform the flexible and complex operations required for efficient recovery of the connection failure. The devices in the data plane take advantage of the application plane's flexibility and programmability features through the abstractions created by the control plane. For example, network monitoring and failure identification can be made extensively by developing and deploying fault detection applications in the management plan.

2.2 Reinforcement learning (RL)

For machine learning, RL is widely used. RL is not restricted to a particular mechanism. To achieving the reward, the agent should interact with the environment [1], which is a basic idea to be followed. The definition of RL is defined by describing a learning problem, and any approach that can address this issue can be considered an RL method. To reach this aim, the agents that defined have to can communicate with the environment, and to a certain degree of knowledge of the environment, and then make an action that can affect the state via perception [4]. With these basic requirements, RL should have one or more targets to create an environment. Markov Decision Processes (MDPs) are a mathematically idealized version of the RL problem for which precise theoretical claims can be made [8].

3 RL BASED EFFICIENT ROUTING IN SDN

In this section, we summary a brief applying RL based routing in SDN with table 1.

Routing is an essential function of a network. In SDN, the flow table within switches is adjusted by the controller to manage the routing. For example, the controller can direct switches to reject flows or route it via a particular path. Ineffective routing choices can lead to network links being overload, and the end-to-end delivery delay increased, which influences SDN's performance. Thus, a principal research issue is how to the routing of traffic flows. Most SDN routing algorithms still use the shortest path algorithm based on the number of hops to provide the optimized route. However, these algorithms do not consider the bandwidth, delay, resource utilization rate, time, and other factors. Besides, routing issues of optimizing can be called as a decision-making task. Thus, reinforcement learning is a practical approach. In the paper [3], the authors proposed a routing QoS security method based on the Q-Learning algorithm and the business flow attribute. Towards this end, this paper built four modules (link discovery, link classification, intensive learning, and Q - value table sending) to assign different paths to different attribute data streams that deployed on the Control layer. Learning conducted by Q-learning and elements for learning are obtained using the link discovery module and link classification module. A total of four factors used when choosing the optimum route, such as link delay, possible bandwidth, packet loss rate, and bandwidth usage. Also, considering the negative and constructive correlation of each factor to find reward and package attributes classified into four categories (interactive services, online media services, interactive services, and data layer services). Depending on the properties of the packet, the weight of each element is obtained, and then the reward can be obtained using element values and weights. Based on this, every time the learning is conducted, it updates the value of the Q table. After the study is completed, the Q value table can easily find the way through the matrix obtained by the properties of the packet through all-new OpenFlow switches. Furthermore, the network resource consumption ratio rationally allocated to avoid local congestion, and the QoS demand of users met as far as possible. However, this approach is based on static topology, while the network model is always dynamic.

Additionally, the paper9 proposed an approach, which combines a deep neural network (DNN) with RL. The work [12] investigated the problem of controller synchronization with a limited budget of sync in distributed SDN, for which their goal was to find the policy that maximizes the benefits of controller synchronizations over time. In a distributed SDN environment, timing is necessary because all controllers have to maintain the same global view. However, in the traditional SDN environment, the lack of consistency between controllers caused severe performance degradation as the network grew. Thus, this paper formulated the problems of controller synchronization as issues with MDP. To solve the MDP problem, they have suggested an RL-based algorithm that uses the DNN to represents its value function, called the DQ Scheduler. DQ scheduler makes optimal decisions about synchronization between controllers through enhanced learning to reduce the degradation of such efficiency. Inter-controller synchronization issues mainly occur during cross-domain routing. Many parts do not have the benefit of synchronization when deciding the direction between domains by the fluid nature of the network, because they are not instantly synchronized. Thus, the Average Path Cost (APC) estimated at the path's average cost is used to quantify the performance of the optimized routing path in the chosen domain route when learning to optimize the synchronization benefits. In other words, learning reinforcement is used to lower APC. To this end, the learning model state is the time slot since the last synchronization, and action is whether or not to synchronize with the chosen domain. Then, rewards become APC Reduction rewards, and learning is Q-Learning. In summary, evaluation results showed that DQ Scheduler provides almost performance higher more than the current SDN controller synchronization solutions. DQ Scheduler is much more flexible in that there are no assumptions on the network, and the policy learning process is automated given any network conditions.

Besides, the work [13] also proposed a routing algorithm based on a semi-supervised learning method which combines the perceptual ability of deep learning with the decision ability of RL. This paper presented an overview of SDN - enabled ultra-dense networking (UDN) crowd management in smart cities with MEC, and also describes an intelligent DRL-based routing solution. They described the problem in terms of balancing service access speed, and the use of network resources due to the high cost of construction and maintenance of the infrastructure. To this end, this paper proposed the DRLS routing algorithm to intelligently manage multiple crowd distribution requests for services across various smart city sectors. The DRL device deployed in the SDN controller is responsible for carrying out the DRLS that is suggested. The state vector of the proposed DRL system, which consists of the location of the request created for service M by the edge node N, the action vector "a" consists of the determining factor, and also the reward successfully sent to the edge node where request provides the appropriate service. Link, the amount of resource usage balance calculated by network load distribution, and the delay in transmitting the request data to access the service via the edge network. It consists of three types of resource balance determined by the distribution of network load and the latency of request data transmission to the edge network to access the service. As a result, compared to the existing OSPF strategy and the improved EOSPF strategy, the DRLS analysis algorithm has achieved much better performance. Although EOSPF puts some effort into the understanding of the dynamic environment, it is limited in its ability to adapt to these changes. However, through the training process, our DRLS will actively learn the evolving network status and smartly pick valid paths for various service requests in the complex environment.

Moreover, AI techniques are growing significantly nowadays, with the ability to analyze some new data samples and to make accurate decisions in a more complicated environment. Consequently, many researchers in the field of network design are now paying attention to using network architecture for AI applications. In 2017, Albert et al.,[6] published their work that is Knowledge-Defined Networking (KDN), which brings huge benefits to routing in SDN but also brings fundamental challenges that need addressing. Thank the KDN [9], Sun, Penghao, and et al., proposed TIDE, smart network control architecture based on DRL [5]. The intensity, diversity, and complexity of the spatio-temporal distribution of traffic flows have significantly increased due to the rapid growth of the communication network, resulting in resolving the network traffic imbalance through routing optimization using QoS. This paper proposed the TIDE network control architecture, a smart network control architecture based on deep reinforcement learning that, without human experience, can dynamically automate routing strategies within an SDN network. The smart agent of AI Plane consisted of the "Collection Decision-Adjust" loop "Collection" step of TIDE is the information collection of network status and performance of routing strategies. The "Decision" stage is an intelligent AI plane algorithm, and the "Adjust" stage is converted to routing policy and takes the form of each switch's flow table party. Through combining RNN-based DDPG with TIDE's routing optimization program, the agent will more effectively view the network status through choosing traffic sequences instead of the immediate network traffic distribution. The state is the order of information regarding network status. Action is the weight matrix specifying the weight of the connection for each dual connection between the routers, and Rewards is the quality of service defined by the QoS parameter. RNN (TIDE relation weight) gives the final result. Thus optimal routing strategies can be dynamically generated for trained networks. This approach not only has tremendous potential to replace traditional routing strategies but also Its architecture ideal for stable environments such as networks and data centres. However, the intelligent agent uses a neural network that takes the view of the network as the input so that it can not be versatile with the dynamic network.

Also, designing a routing method that uses DRL to achieve standardized and configurable optimization in continuous times is a big challenge. To address this issue, this paper [11] proposed a machine learning-based SDN framework, which uses a novel DRL mechanism called DDPG [5] to optimize routing in SDN. As network traffic grows exponentially, there is a need to streamline the SDN routing process while maintaining the quality of service. By adding an intelligent decision module of ML to the control plane, a global network strategy was created by the controller based on approach. Deep Deuteronomic Policy Gradient (DDPG), which incorporates the DPG method in an actor-critic framework with the DQN method, using neural networks produces strategic functions and Q functions and provides a practical and stable model of discrete controlled action. A routing optimization method called DROM, based on DDPG, has proposed the realization of social, real-time, and personalized network intelligence control and management in continuous time. The DROM agent can change the path of the data flow by changing the traffic matrix (TM) of the network load and action by changing the link weight of the network. The agent's Rewards have the advantage of automatically optimizing custom performance parameters such as delay, forwarding path length, and throughput associated with network operation and maintenance strategies, thereby continuously realizing real-time network control. The result of the experiment indicates DROM has good convergence and productivity, and DROM will boost network performance with reliable and superior routing services compared to existing routing solutions.

What is more, to improve the efficiency of QoS-aware routing, in [7], the authors exploited a DRL agent with convolutionary neural networks in KDN [6] context. The QoS aware routing problem is challenging when several flows coexist within the same system. To this end, the Knowledge Plane is inserted through the control loop to the SDN paradigm to provide integration, optimization, testing, and estimation. The proposed mechanism will learn the mutual impacts between flows in the systems by implementing convolutional neural networks; hence, it can provide better routing configurations. The recommended DRL agent for convolution is defined as State as TM, Action as a link-weight vector and reward as the average, and qualified flow average of the QoS matrix. This agent's goal is to confirm the optimal policy for mapping from state space to action space by repeatedly enhancing the knowledge of the relation between state, behavior, and compensation for actors and essential neural depth. A vanilla solution with the DDPG algorithm 7 was used to solve the routing problem. In heterogeneous power networks, the difference even expanded, showing the benefits of convolutional neural networks in dealing with complex scenarios. Knowledge Defined Networking is a possible data network model for the future. In the form of routing this paper studied neural network architectures in deep reinforcement learning (DRL). The proposed mechanism can learn the mutual impacts between flows in the networks by implementing convolutionary neural networks; hence it can provide better routing configurations. The differences in heterogeneous ability networks are also broadening, suggesting the benefits of convolutionary neural networks in dealing with complex scenarios.

4 CONCLUSION

In this paper, we provide an overview of applying RL based efficient Routing in SDN. This paper described the necessary background of SDN and a brief of RL. Then we summarized the recent research contributions to provide more intelligent network behavior in the routing SDN approach. The mechanisms combined RL with DNNs have been applied in different scopes, such as controller synchronization and routing optimization. Employing RL techniques in SDN routing showed an efficient to enforce routing to the dynamic network, so they can provide excellent levels of QoS while optimizing resource utilization. In summary, this paper aims to discuss briefly how RL algorithms operate and how they are used to solve SDN problems. We expected our discussion and discovery would open up a new avenue for SDN growth and the introduction of a smarter network. This paper makes an effort to motivate the readers, discover the application of RL-based routing in SDN, and exploit potential benefits for network performance optimization goals. That also is research in the future, and we aim to study and implement variants of RL techniques in order to optimize the performance of SDN networks. We will focus on studying DRL, GNN (Graph Neural Network) for maximizing throughput and minimizing delay issues.

ACKNOWLEDGMENTS

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning (NRF-2017R1A2B40 12559). This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the ITRC(Information Technology Research Center) support program(IITP-2020-2016-0-00314) supervised by the IITP(Institute for Information & communications Technology Planning & Evaluation).

REFERENCES

[1] 2005. Reinforcement Learning: An Introduction. *IEEE Transactions on Neural Networks* 16, 1 (2005), 285–286.

[2] Open Networking Fundation. 2012. Software-defined networking: The new norm for networks. *ONF White Paper* 2, 2-6 (2012), 11.

[3] Zijin Jin, Weifei Zang, Yiming Jiang, and Julong Lan. 2018. A QLearning Based Business Differentiating Routing Mechanism in SDN Architecture. In *Journal of Physics: Conference Series*, Vol. 1168.

[4] Frank L Lewis and Draguna Vrabie. 2009. Reinforcement learning and adaptive dynamic programming for feedback control. *IEEE circuits and systems magazine* 9, 3 (2009), 32–50.

[5] Timothy Paul Lillicrap, Jonathan James Hunt, Alexander Pritzel, Nicolas Manfred Otto Heess, Tom Erez, Yuval Tassa, David Silver, and Daniel Pieter Wierstra. 2017. Continuous control with deep reinforcement learning. (Jan. 26 2017). US Patent App. 15/217,758.

[6] Albert Mestres, Alberto Rodriguez-Natal, Josep Carner, Pere Barlet-Ros, Eduard Alarcón, Marc Solé, Victor Muntés-Mulero, David Meyer, Sharon Barkai, Mike J Hibbett, et al. 2017. Knowledge-defined networking. *ACM SIGCOMM Computer Communication Review* 47, 3 (2017), 2–10.

[7] Tran Anh Quang Pham, Yassine Hadjadj-Aoul, and Abdelkader Outtagarts. 2018. Deep reinforcement learning based qos-aware routing in knowledge-defined networking. In *International Conference on Heterogeneous Networking for Quality, Reliability, Security and Robustness*. Springer, 14–26.

[8] Martin L Puterman. 2014. *Markov decision processes: discrete stochastic dynamic programming*. John Wiley & Sons.

[9] Penghao Sun, Yuxiang Hu, Julong Lan, Le Tian, and Min Chen. 2019. TIDE: Time-relevant deep reinforcement learning for routing optimization. *Future Generation Computer Systems* 99 (2019), 401–409.

[10] Junfeng Xie, F Richard Yu, Tao Huang, Renchao Xie, Jiang Liu, Chenmeng Wang, and Yunjie Liu. 2018. A survey of machine learning techniques applied to software defined networking (SDN): Research issues and challenges. *IEEE Communications Surveys & Tutorials* 21, 1 (2018), 393–430.

[11] Changhe Yu, Julong Lan, Zehua Guo, and Yuxiang Hu. 2018. DROM: Optimizing the routing in software-defined networks with deep reinforcement learning. *IEEE Access* 6 (2018), 64533–64539.

[12] Ziyao Zhang, Liang Ma, Konstantinos Poularakis, Kin K Leung, and Lingfei Wu. 2019. DQ scheduler: Deep reinforcement learning based controller synchronization in distributed SDN. In *ICC 2019-2019 IEEE International Conference on Communications (ICC)*. IEEE, 1–7.

[13] Lei Zhao, Jiadai Wang, Jiajia Liu, and Nei Kato. 2019. Routing for crowd management in smart cities: A deep reinforcement learning perspective. *IEEE Communications Magazine* 57, 4 (2019), 88–93.