Guest Editorial
Leveraging Machine Learning in SDN/NFV-Based Networks

David S. L. Wei, *Senior Member, IEEE*, Kaiping Xue, *Senior Member, IEEE*,
Roberto Bruschi, *Senior Member, IEEE*, and Stefan Schmid

I. INTRODUCTION

A KEY trend of current network evolution is in the direction of network softwarization and virtualization. These technological paradigms aim to enable a network to be programmable in a way that makes the network more flexible, scalable, and reliable, and in turn leads to agile service deployment and lower capital and operational expenses. So far, two related widely adopted solutions are software defined networks (SDN) and network function virtualization (NFV). There is one main difference between these two new networking paradigms. SDN separates the control plane from the data plane through a well-defined programming interface, such that the centralized controller can have a complete view of the entire network, while NFV decouples network functions from dedicated physical equipment by means of virtualization technology, and runs the virtual network functions (VNFs) in the general purpose physical or virtual network appliances. Both approaches make the network programmable in order to have the aforementioned desired features. SDN and NFV do not depend on each other, and they actually complement each other. They can work well individually and can also work in tandem for performance reasons. Due to such advantages, both SDN and NFV have become key enabling technologies for 5G networks, and have also been used in a wide range of important areas including IoT, mobile edge computing, smart grid, cloud datacenters, and cognition-based networks.

Although SDN and NFV facilitate the flexibility and scalability of network services and make the deployment of network services faster and cheaper, such software-based solution also introduces new problems, including throughput performance degradation and unstable jitter. More specifically, in SDN/NFV-based networks, resource management, traffic control, and network security are among the challenges that telecommunication service providers must overcome in order to provide better services to users and increase their revenue. Meanwhile, machine learning (ML) has achieved great success in solving problems in various areas ranging from natural language processing and voice recognition to autopilot and strategic game playing. It is believed that ML also has high potentiality in addressing the aforementioned challenges in SDN/NFV-based networks, especially in the elastic deployment of virtual network functions (VNFs), dynamic service provisioning, adaptive traffic control, and security, as ML is a technology that can effectively extract knowledge from data and accurately predict the future resource requirements of each virtualized software-based appliance and the future service demands of each user. Researchers from both academia and industries have started their research on exploring various ML techniques to help address those key issues in SDN/NFV-based networks. The objective of this special issue is to feature recent research in this direction.

II. SUMMARY OF THE ARTICLES IN THIS SI

Though research in this direction is still in its nascent stage, this special issue still attracts 35 high-quality submissions from around the world. Due to limited slots, only the following 11 articles are selected for publication after a rigorous review process. The achievements of these works are validated via analysis, simulation, experimentation, and system implementation. They cover the topics of VNF and service placement optimization, resource allocation in network slicing, OpenFlow switches, and congestion window control in 5G networks. We summarize these articles in the order of the listed topics in what follows.

A. VNF and Resource Placement Optimization

The first six articles address the issue of VNF or resource placement and chaining in NFV. In the article "Optimizing NFV Chain Deployment in Software-defined Cellular Core", J. Zhang *et al.* address the problem of optimizing VNF chaining in a software-defined cellular core. The authors propose a two-stage optimization framework that first minimizes the service chain deployment cost, and then determines which

D. S. L. Wei is with the Department of Computer and Information Science, Fordham University, New York, NY 10458 USA (e-mail: wei@cis.fordham.edu).

K. Xue is with the School of Cyber Security, University of Science and Technology of China, Hefei 230027, China, and also with the Department of Electronic Engineering and Information Science, University of Science and Technology of China, Hefei 230027, China (e-mail: kpxue@ustc.edu.cn).

R. Bruschi is with the National Laboratory of Smart, Sustainable and Secure Internet Technologies and Infrastructures (S3ITI), Italian National Consortium for Telecommunications (CNIT), 16145 Genoa, Italy (e-mail: roberto.bruschi@unige.it).

S. Schmid is with the Faculty of Computer Science, University of Vienna, 1010 Vienna, Austria (e-mail: schmiste@gmail.com).

Digital Object Identifier 10.1109/JSAC.2019.2959197

VNF should be deployed onto which CPU core to balance the CPU processing capability. Their simulations show that their proposed scheme can significantly reduce the capital cost and also increase the throughput. In the article "Optimal VNF Placement via Deep Reinforcement Learning in SDN/NFV-Enabled Networks", J. Pei *et al.* address VNF placement problem under dynamic network loading in SDN/NFV-enabled networks. The authors formulate the problem as a binary integer programming model aiming at minimizing the total cost consisting of VNF placement cost and VNFI running cost. They then propose a scheme to solve this problem using deep reinforcement learning. Evaluation results show that their scheme can help improve network performance in terms of various performance metrics. In the article "Intelligent VNF Orchestration and Flow Scheduling via Model-assisted Deep Reinforcement Learning", L. Gu *et al.* investigate the problem of orchestrating VNFs via VNF activation and deactivation to maximize the overall network utility. The authors design a model-assisted deep reinforcement learning (DRL) framework for accelerating the training in VNF orchestration. Experimental results validate the high efficiency of their model-assisted DRL framework as it not only converges much faster than the traditional DRL algorithm but it also achieves higher performance. In the article "Virtual Network Function placement optimization with Deep Reinforcement Learning", R. Solozabal *et al.* solve the VNF-FGE problem by means of a Reinforcement Learning approach to model a placement policy in a Network Function Virtualization infrastructure. The authors extend the Neural Combinatorial Optimization theory by considering constraints in the definition of the problem so that the resulting agent is able to learn placement decisions by exploring the NFV infrastructure with the aim of minimizing the overall power consumption. In the article "Management and Orchestration of Virtual Network Functions via Deep Reinforcement Learning", J. S. Pujol *et al.* present a novel DRL algorithm for autonomous management and orchestration of VNFs, where the CU learns to re-configure resources, i.e., CPU and memory, to deploy new VNF instances or to offload VNFs to a central cloud. They show the performance in numerical results and map the results to 5G key performance indicators. In the article "A Dynamic Reliability-Aware Service Placement for Network Function Virtualization (NFV)", M. K. Farshbafan *et al.* investigate a reliability-aware service placement by taking the dynamic nature of the service arrivals and departures into consideration. They adopt a model based on an infinite horizon Markov decision process for dynamic reliability-aware service placement with the consideration of simultaneous allocation of the main and backup servers. The superiority of their proposed model is verified via extensive simulations.

B. Network Slicing

The next three articles are about resource management in network slicing. In the article "GAN-powered Deep Distributional Reinforcement Learning for Resource Management in Network Slicing", Y. Hua *et al.* propose a generative adversarial network-powered deep distributional Q network to learn the optimum solution for demand-aware resource management in network slicing. They devise a reward-clipping mechanism to stabilize their scheme's training against the effects of widely-spanning utility values, and also develop another dueling scheme. They then verify the superiority of the two proposed algorithms through extensive simulations. In the article "Offline SLA-Constrained Deep Learning for 5G Networks Reliable and Dynamic End-to-End Slicing", H. Chergui *et al.* address the issue of resource provisioning as an enabler for end-to-end dynamic slicing in SDN/NFV-based 5G networks. They first present a low-complexity network slices' traffic predictor and then use the predicted traffic to feed several deep learning models trained offline to perform end-to-end dynamic and reliable resource slicing under dataset-dependent generalized non-convex SLA constraints. In the article "DeepCog: Optimizing Resource Provisioning in Network Slicing with AI-based Capacity Forecasting", D. Bega *et al.* present a data analytics tool, named DeepCog, for the cognitive management of resources in sliced 5G networks. DeepCog is a deep neural network architecture inspired by advances in image processing and trained via a dedicated loss function. Empirical evaluations with real-world metropolitan-scale data demonstrate the effectiveness of their proposed solution.

C. OpenFlow Switches

The next article is about OpenFlow switches. In the article "STEREOS: Smart Table EntRy Eviction for OpenFlow Switches", H. Yang *et al.* propose machine learning techniques to optimize flow entry eviction in OpenFlow switches. They discuss implementation issues, including model selection, model sizing, overhead, and feature quantization. Network-level simulations demonstrate that their techniques can greatly reduce control overhead, increase network throughput, and reduce packet loss rate.

D. Congestion Window Control in 5G MEC

The next article is about congestion window control in 5G mobile edge computing. In the article "Adaptive Online Decision Method for Initial Congestion Window in 5G Mobile Edge Computing using Deep Reinforcement Learning", R. Xie *et al.* investigate the initial congestion window (IW) decision problem in mobile edge computing. The objective is to adaptively adjust initial congestion window such that flow completion time is optimized while congestion is minimized. They propose an adaptive online decision method to solve the problem. Their simulations using a 5G mmWave MEC simulator demonstrate that their algorithm can effectively reduce flow completion time with little congestion and can adapt IW to dynamic network conditions.

III. CONCLUSION AND ACKNOWLEDGEMENT

In conclusion, the authors sincerely hope that this special issue provides up-to-date and valuable research information for the researchers currently conducting research in network softwarization. Taking this opportunity, the authors deeply appreciate Dr. Moshe Zukerman who rendered prompt advice and assistance in the preparation of this special issue.

Special thanks go to Dr. Raouf Boutaba for his advice and encouragement in the initial stage of their proposal. The authors also thank Janine Bruttin for her quick responses to our various queries so that this special issue could be timely processed. The authors are also grateful to the authors who submitted their quality works to this special issue and thank all reviewers for their efforts and valuable reviews.

David S. L. Wei (SM'07) received the Ph.D. degree in computer and information science from the University of Pennsylvania in 1991. From May 1993 to August 1997, he was with the Faculty of Computer Science and Engineering, University of Aizu, Japan, as an Associate Professor and then a Full Professor. He is currently a Full Professor with the Computer and Information Science Department, Fordham University. He has authored or coauthored more than 120 technical articles in the areas of parallel and distributed processing, wireless networks and mobile computing, optical networks, P2P communications, big data, cloud computing, and IoT in various archival journals and conference proceedings. He currently focuses his research efforts on cloud and edge computing, IoT, 5G, big data, and machine learning. He served on the program committee and was a Session Chair for several reputed international conferences. He is presently an Associate Editor of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS for the Series on Network Softwarization & Enablers, and ever served as an Associate Editor and a Lead Guest Editor for some leading journals, such as IEEE J-SAC, IEEE TCC, and IEEE TBD.

Kaiping Xue (M'09–SM'15) received the bachelor's degree from the Department of Information Security, University of Science and Technology of China (USTC), in 2003, and the Ph.D. degree from the Department of Electronic Engineering and Information Science (EEIS), USTC, in 2007. From May 2012 to May 2013, he was a Postdoctoral Researcher with the Department of Electrical and Computer Engineering, University of Florida. He is currently an Associate Professor with the School of Cyber Security and the Department of EEIS, USTC. He has authored or coauthored more than 100 technical articles in the areas of communications, networks, and network security. His research interests include next-generation Internet, distributed networks, and network security. He is an IET fellow. He was a recipient of the Best Paper Award at IEEE MSN 2017 and the Best Paper Runner-Up Award at IEEE MASS 2018. He serves on the Editorial Board of several journals, including the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS (TWC), the IEEE TRANSACTIONS ON NETWORK AND SERVICE MANAGEMENT (TNSM), *Ad Hoc Networks*, IEEE ACCESS, and *China Communications*. He has also served as a Guest Editor of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS (JSAC) and a Lead Guest Editor of the IEEE Communications Magazine.

Roberto Bruschi (M'08–SM'13) received the Ph.D. degree in electronic engineering from the University of Genoa in 2006. From 2010 to 2019, he has been a Researcher with the S2N National Laboratory, National Inter-University Consortium for Telecommunications. Since December 2019, he has been an Associate Professor with the University of Genoa. He has coauthored over 115 articles in international journals, book chapters, and conference proceedings. His current main research interests include green networking, 5G, NFV, SDN, and edge computing. He was a recipient of the Best Paper Award at the IEEE ICC 2009 and the 2010 IEEE GreenCom Workshop, and the Best Demonstration Award at IEEE/IFIP IM 2017. He has been the Coordinator of the European H2020 INPUT project and of the Italian FiR GreenNet project. He also served as a Project Manager of the European FP7 ECONET Project and a WP Leader of the H2020 5G-PPP MATILDA project. He is an Associate Editor for the IEEE TRANSACTIONS ON GREEN COMMUNICATIONS AND NETWORKING and the IEEE Communications Magazine.

Stefan Schmid received the Diploma (M.Sc.) degree in computer science from ETH Zurich, Switzerland (minor: micro/macro economics, internship: CERN) and the Ph.D. degree from the Distributed Computing Group, ETH Zurich, led by Prof. Roger Wattenhofer. As a postdoc, he worked with Prof. Christian Scheideler at the Chair for Efficient Algorithms, Technical University of Munich, and at the Chair for Theory of Distributed Systems, University of Paderborn, Germany. From 2009 to 2015, he was a Senior Research Scientist with the Telekom Innovation Laboratories (T-Labs) and with the Internet Network Architectures Group, TU Berlin, Germany, headed by Prof. Anja Feldmann. In 2013/14, he was an INP Visiting Professor with CNRS (LAAS), Toulouse, France, and in 2014, a Visiting Professor with the Université catholique de Louvain (UCL), Louvain-la-Neuve, Belgium. From 2015 to 2017, he was a tenured Associate Professor with the Distributed, Embedded and Intelligent Systems Group, Aalborg University, Denmark, and continued working part-time at TU Berlin, Germany. He is currently a Professor with the Faculty of Computer Science, University of Vienna, Austria. His research interests revolve around the fundamental and algorithmic problems of networked and distributed systems. He received the IEEE Communications Society ITC Early Career Award 2016. Since 2015, he has been the Editor of the Distributed Computing Column of the *Bulletin of the European Association of Theoretical Computer Science* (BEATCS), since 2016, as an Associate Editor of the IEEE TRANSACTIONS ON NETWORK AND SERVICE MANAGEMENT (TNSM), and since 2019, as an Editor of the IEEE/ACM TRANSACTIONS ON NETWORKING (ToN).