Series Editorial: Inauguration Issue of the Series on Machine Learning in Communications and Networks

Geoffrey Y. Li, *Fellow, IEEE*, Walid Saad, *Fellow, IEEE*, Ayfer Ozgur, Peter Kairouz, Zhijin Qin, Jakob Hoydis, Zhu Han, *Fellow, IEEE*, Deniz Gunduz, and Jaafar Elmirghani

I. Introduction

In the era of the new generation of communication systems, data traffic is expected to continuously strain the capacity of future communication networks. Along with the remarkable growth in data traffic, new applications, such as wearable devices, autonomous systems, and the Internet of Things (IoT), continue to emerge and generate even more data traffic with vastly different requirements. This growth in the application domain brings forward an inevitable need for more intelligent processing, operation, and optimization of future communication networks.

With the successes of machine learning, especially deep learning, in various applications, researchers in the communications community have also applied it to improve the communications system performance or make networks more intelligent. Particularly, in intelligent communications, enormous efforts are being made to improve the system performance in physical layer signal processing, resource allocation, and network design.

The inauguration issue of the "Series on Machine Learning in Communications and Networks" has received a plethora of high-quality papers dealing with diverse aspects and topics of intelligent communication systems, ranging from intelligent signal processing, resource allocation, and network design to various wireless application scenarios. The Call for Papers received an especially strong response from the community. We could only accept 32 original contributions with 20 ones in the inauguration issue and 12 ones in the next issue to be published in July 2021. In the following, we provide a brief review of the key contributions of this issue according to the topics.

Geoffrey Y. Li and Deniz Gunduz are with the Department of Electrical and Electronic Engineering, Imperial College London, London SW7 2BT, U.K.

Walid Saad is with the Electrical and Computer Engineering Department, Virginia Tech, Blacksburg, VA 24060 USA.

Ayfer Ozgur is with the Department of Electrical Engineering, Stanford University, Stanford, CA 94305 USA.

Peter Kairouz is with Google AI, Mountain View, CA 94043 USA.

Zhijin Qin is with the School of Electronic Engineering and Computer Science, Queen Mary University of London, London E1 4NS, U.K.

Jakob Hoydis is with Nokia Bell Labs, Paris-Saclay, 91620 Nozay, France.

Zhu Han is with the Electrical and Computer Engineering Department, University of Houston, Houston, TX 77004 USA.

Jaafar Elmirghani is with the School of Electronic and Electrical Engineering, University of Leeds, Leeds LS2 9JT, U.K.

Digital Object Identifier 10.1109/JSAC.2020.3036785

II. Signal Processing

This issue consists of five papers that deal with various problems in signal processing, including detection, estimation, prediction, and compression. We have three papers related to channel estimation and predication. The paper, titled "Solving sparse linear inverse problems in communication systems: A deep learning approach with adaptive depth," by Chen *et al.*, investigates the sparse signal recovery problem. The paper proposed an end-to-end deep learning structure with adaptive depth, which could be adjusted for different tasks. The developed structure is then applied to random access and channel estimation in massive multi-in-multi-out (MIMO) systems to demonstrate its effectiveness. "High dimensional channel estimation using deep generative networks" by Balevi *et al.* adopts compressed sensing for the sparse signal recovery problem in channel estimation for millimeter waves. With the proposed generative networks, the channel estimation requires fewer pilots and eliminates the requirements on a sparsifying basis without performance degradation. "Model refinement learning and an example on channel estimation with universal noise model" by Wang *et al.* proposes a scheme to deal with the channel estimation with unknown noise. A universal mixture of the Gaussian model for the unknown distribution of noise is utilized to update the parameters to fit the observed data. Moreover, an online channel estimation algorithm with linear computational complexity is proposed. Simulation results reveal that the proposed design outperforms both the model-based and data-based counterparts.

The paper, titled "Generative-adversarial-network enabled signal detection for communication systems with unknown channel models," by Sun *et al.*, deals with the signal detection without channel state information in the Viterbi algorithm. By extending ViterbiNet (Shlezinger *et al.*), a novel architecture using generative adversarial networks (GAN) is proposed to learn the channel transition probability (CTP) to make the classic Viterbi algorithm implementable. The fine-tuning of the learned CTP is achieved so that the proposed design is adaptive to dynamic channel conditions without requiring training from scratch. The paper, titled "Deep multi-task learning for cooperative NOMA: System design and principles," by Cheng *et al.*, considers the channel coding and power allocation for the nonorthogonal multiple access (NOMA) system. A novel hybrid-cascaded deep neural network (DNN) architecture is proposed to optimize the entire system in terms of the bit-error rate (BER). Simulation results demonstrate its advantages over the benchmarks in various scenarios.

III. Learn to Transmit and Receive

There are two papers in the category of learning to transmit and receive. The paper, titled "Perm2vec: Attentive graph permutation selection for decoding of error correction codes," by A. Caciularu *et al.*, leverages neural transformer networks to predict the permutation of a noisy codeword that maximizes the probability of leading to a correctly decoded codeword. The effectiveness of the proposed scheme is evaluated for belief propagation (BP) decoding of several short BCH codes. The paper, titled "Wireless image retrieval at the edge" by M. Jankowski *et al.*, investigates the problem of image retrieval, where a sensing device captures an image and transmits a related low-dimensional signature over a noisy channel to a server that tries to retrieve similar images from a database. The authors propose a novel neural network architecture that directly transforms the observed image into channel symbols for transmission and demonstrates various benefits of this approach over alternative solutions.

IV. Resource Management and Network Optimization

Three papers in this issue address resource allocation using machine learning. In the paper, titled "Graph neural networks for scalable radio resource management: Architecture design and theoretical analysis," by Shen *et al.*, graph neural networks (GNNs) are proposed to solve large-scale radio resource management problems. A class of neural networks, named message passing graph neural networks (MPGNNs), are demonstrated not only to satisfy the permutation equivalence property but also to be able to generalize to large-scale problems while enjoying high computational efficiency. The paper, titled "Situation-aware resource allocation for multi-dimensional intelligent multiple access: A proactive deep learning framework," by Liu *et al.*, considers a time-division-duplex downlink cellular scenario. A deep learning-based framework for a multidimensional intelligent multiple access (MD-IMA) scheme is developed for beyond 5G wireless networks to meet the real-time and diverse QoS requirements by fully utilizing the available radio resources in heterogeneous domains. To achieve intelligent operation of MD-IMA, the proposed deep learning scheme is a convergence of long short-term memory (LSTM) and deep reinforcement learning (DRL). "Multi-agent reinforcement learning based resource management in MEC- and UAV-assisted vehicular networks," by Peng *et al.*, considers the multidimensional resource management problem for unmanned aerial vehicle (UAV)-assisted vehicular networks with mobile edge computing servers mounted on both the marco eNodeBs and UAV. A multiagent deep deterministic policy gradient (MADDPG)-based method is developed to cooperatively make association decisions and allocate proper amounts of resources to vehicles to maximize the number of offloaded tasks while satisfying their heterogeneous quality-of-service requirements.

V. Semantic Communications

The paper, titled "A lite distributed semantic communication system for Internet of Things," by Xie *et al.*, investigates the semantic communication system, which is treated as the second level of communications. Instead of trying to recover bit sequences successfully, semantic communications aims to realize successful semantic information exchange. In this paper, a distributed semantic communication system is proposed for text transmission with low complexity, where the data transmission from the IoT devices to the cloud/edge works at the semantic level to improve transmission efficiency. Moreover, by pruning the model redundancy and lowering the weight resolution, the proposed scheme becomes affordable for IoT devices and the bandwidth required for model weight transmission between IoT devices and the cloud/edge is reduced significantly.

VI. Distributed/Federated Learning and Communications

We have two papers that focus on the privacy and security aspects of federated learning. Privacy of local datasets is studied in "Privacy for free: Wireless federated learning via uncoded transmission with adaptive power control," by Liu *et al.*, which argues that the additive noise in the channel can be exploited to provide natural privacy protection against information leakage from gradients, when over-the-air computation is exploited for federated learning in wireless systems. Privacy of user data is also considered in "Optimal contract design for efficient federated learning with multi-dimensional private information," by Ding *et al.*, but the authors focus on the incentive mechanisms encouraging users' participation in federated learning. They propose a multidimensional contract theoretic approach, which they simplify by summarizing the user's private information into a 1-D criterion.

We have three papers that focus on the communication bottleneck in federated learning and propose various approaches to speed it up. In "Fast-convergent federated learning", Nguyen *et al.* propose a sampling algorithm to select the devices that will participate in each round, taking into account the gradient information. They show that the proposed algorithm has a better convergence speed compared to the state-of-the-art. Another paper aiming to speed up federated learning is "Accelerating DNN training in wireless federated edge learning systems" by Ren *et al.* It proposes a new performance evaluation criterion, called learning efficiency. Then, for both the CPU and GPU scenarios, the authors formulate a training acceleration optimization problem to solve for the optimal joint batch size selection and communication resource allocation. Federated learning can suffer from slow convergence due to heterogeneity and randomness in computing power and communication link qualities across devices. In the paper, "Coded computing for low-latency federated learning over wireless edge networks," by Prakash *et al.*, the authors address this issue by employing coding, where the devices share a coded version of their local data sets with the parameter server, which is then used to compensate for straggling servers.

VII. Selected Topics

This issue also contains three papers on edge service, TCP scheme, and optical communications. The paper "Seek common while shelving differences: Orchestrating deep neural networks for edge service provisioning," by Chen *et al.*, considers the service provisioning problem in edge computing systems. The authors propose a novel framework based on multi-agent deep reinforcement learning, which employs a distributed neural network orchestration scheme as well as knowledge distillation. The proposed approach coordinates edge servers to maximize the application service provider's reward in a fully distributed manner. For the paper "Wanna Make Your TCP Scheme Great for Cellular Networks? Let Machines Do It for You!" by Abbasloo *et al.*, to improve the performance of TCP, it introduces a deep reinforcement learning plug-in that atomically steers throughput-oriented TCP algorithms toward achieving the desired application performance. A range of new and old TCP schemes are considered showing that significant performance improvements can be achieved through the use of the proposed plug-in. The paper "Physics-based deep learning for fiber-optic communication systems," by Häger *et al.*, develops a novel approach that integrates machine-learning with the domain knowledge from fiber-optic communication systems to efficiently solve the nonlinear Schrödinger equation (NLSE). The proposed approach explores the similarities between the popular split-step method (SSM) for numerically solving the NLSE and a DNN. The authors use this connection to build a physics-based machine-learning model, which has several advantages over solely using neural networks as a "black-box" function approximator. The authors show that the proposed method can be used effectively for low-complexity nonlinear equalization where the task is to efficiently invert the NLSE (i.e., digital back-propagation). The results show that the proposed approach can reduce the equalization complexity by orders of magnitude in comparison with baseline schemes.

Acknowledgment

We would like to express our great gratitude to all authors for their submissions and over 400 anonymous reviewers for their insightful reviews and suggestions that have helped maintain the high quality of the Series.

We are also grateful to those experts, including Elisabeth de Carvalho, Alexios Balatsoukas-Stimming, Xiangwei Zhou, Tommy Svensson, and Khaled B. Letaief, who have foreseen the importance of machine learning for communications and put forward the proposal for the IEEE Journal on Selected Areas in Communications (JSAC) Series. We would like to thank Professor Robert Schober (IEEE ComSoc VP for publications), Professor Michele Zorzi (IEEE ComSoc Journals Board Director), and Professor Raouf Boutaba (the EiC of the IEEE JSAC) for their support and kind guidance, and Dr. Le Liang (Intel Labs) for helping address various logistic issues.

Finally, we would like to acknowledge our series editorial board for their diligent work during the paper review and decision process. The editorial board of our Series is as the following.

Finally, we wish the contents of our Series will inspire the readers to investigate the challenging and open problems in the field of machine learning in communications.