Intelligent Resource Allocation in Wireless Communications Systems

Woongsup Lee, Ohyun Jo, and Minhoe Kim

The emergence of DL represents a potential paradigm shift in the design of WCS, from conventional handcrafted schemes based on mathematical models with assumptions, to autonomous schemes based on DL using large sets of data. The authors discuss the essential elements of DL and investigate an intelligent RA scheme based on a DNN, in which multiple goals with various constraints can be satisfied through DL.

Abstract

The emergence of DL represents a potential paradigm shift in the design of WCS, from conventional handcrafted schemes based on mathematical models with assumptions, to autonomous schemes based on DL using large sets of data. In this article, we discuss the essential elements of DL and investigate an intelligent RA scheme based on a DNN, in which multiple goals with various constraints can be satisfied through DL. Having confirmed the optimality and feasibility of DNN-based RA through simulation, we discuss some of the key technical challenges that remain problematic for the application of DL in the practical application of WCS.

Introduction

Deep learning (DL) is a technique based on deep neural networks (DNNs), which relates to the activity of neurons in the brain, and which has seen a recent surge in interest. The current enthusiasm for DL is mainly due to its significant advantages in terms of its better performance than conventional schemes[1, 2]. For example, image classification schemes based on DL can achieve far greater accuracy than handcrafted conventional schemes based on analytical models, and have even surpassed human-level performance in some cases [1, 2]. The use of DL is not confined to simple classification; it also shows notable performances in more complicated tasks, such as the understanding of semantic scenes [1].

The advent of DL could cause a change in the paradigm of research, from the design of schemes by careful engineering based on mathematical models, to an approach based on learning, in which the proper scheme is designed autonomously by observing large amounts of data (see Fig. 1). For example, conventional image classification relies on handcrafted complex feature detectors that are engineered and rely on expertise, for example, edge detectors. However, more accurate feature detectors can be derived using DL with large amounts of image data.

Researchers have recently begun to apply DL in many aspects of wireless communications systems (WCS), especially in classification tasks such as traffic and channel estimation. In [3], for example, it was shown that the type of data traffic can be determined accurately using a DNN. Moreover, in [4], DL-based channel estimation and signal detection in orthogonal frequency-division multiplexing (OFDM) systems was considered. DL has also been applied to tasks more complicated than classification, for example, encoder and decoder for optical wireless communications [5] and sparse code multiple access (SCMA) [6], both of which were developed using a DNN-based autoencoder.

One interesting characteristic of DL is that a DNN can be considered to be a universal approximator [7] that is capable of approximating an arbitrary continuous function, allowing it to emulate the behavior of highly nonlinear and complex systems. Moreover, the training of a DNN in an end-to-end manner enables an optimal strategy to be devised, without the need to solve any complicated problems explicitly (see Fig. 1).

In this sense, DL has already been applied to resource allocation (RA) in WCS, and optimization has been used previously. Unlike conventional approaches in which the optimal strategy for RA is derived from the analytical system model using a set of assumptions, an approach based on DL can be used to derive the optimal strategy directly from actual channel data, thanks to its ability to adapt to the environment, leading to a higher performance in practice. In an approach based on DL, the general solver for optimal RA can be derived by finding a strategy with a low computation time even for changing parameters, for example, channel gain [4, 8]. The authors of [7] used a DNN to reconstruct the transmit power of a scheme based on weighted minimum mean square error (WMMSE) in order to solve problems associated with its high computation time. In [8], the transmit power was optimized in order to maximize either the spectral efficiency (SE) or the energy efficiency (EE), where the optimal strategy for transmit power control was achieved through DL without the need to derive an explicit mathematical formulation.

In this article, we focus on intelligent RA in WCS via the use of DL. To this end, we first describe the basic principles of DL. We then discuss DNN architecture, which can be trained to derive a general strategy for RA in order to achieve diverse goals, that is, the maximization of SE and EE, and the minimization of transmit power, while satisfying the constraints on interference and quality-of-service (QoS). Finally, we set out the key research challenges regarding the application of DL, before offering some key conclusions.

Digital Object Identifier:
10.1109/MCOM.001.1900199

Woongsup Lee is with Gyeongsang National University; Ohyun Jo is with Chungbuk National University; Minhoe Kim is with EURECOM.

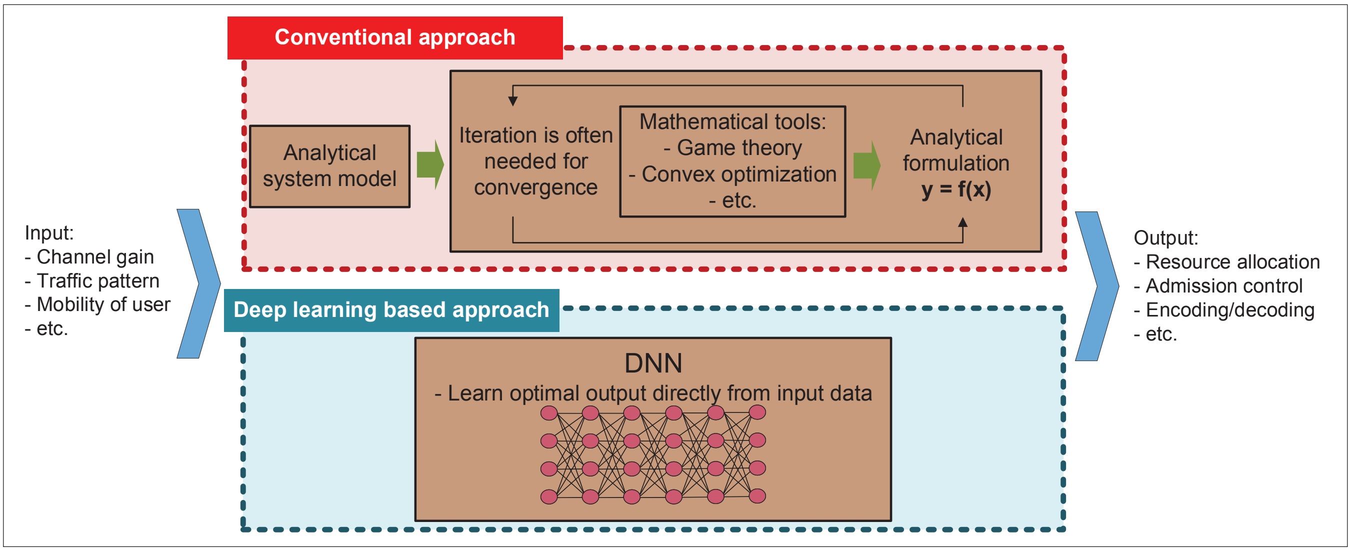


Figure 1. Comparison between deep learning and conventional approaches

Fundamentals of Deep Learning

In this section, we describe the basic components and structure of a Neural Network (NN) and discuss how to train it. Then, we turn our attention to DL and the benefits of the use of DL in WCS.

Structure of Neural Network

NNs comprise DNNs, and the theory behind them is contained in the subcategory of machine learning. NNs are inspired by the operation of the brain, which is composed of neurons that are inter-connected. The output from each neuron can be the input to other neurons, and each neuron activates its output when the inputs satisfy a given condition. In other words, each neuron can be considered to be a biological switch. In an NN, the connection between neurons (nodes) is modeled using matrix multiplication, and the activation of neurons is modeled using nonlinear functions, for example, sigmoid, which provide the ability to model nonlinear characteristics. Accordingly, the NN can be considered to be a collection of computation and activation functions of a matrix, and the calculation of the output from the NN for one set of sample data, that is, inference, usually entails a low computational overhead.

Depending on the existence or not of a feedback loop in the network, the NN can be assigned to either of two categories.

Feedforward Neural Network (FNN): In a FNN, there is no feedback loop, such that the previous input data does not affect the current output, that is, it is 'memoryless'. The FNN is used when the data are not correlated with each other, for example, in the case of image data. In WCS, this type of NN is used for spectrum sensing and the power control of users [7–9].

Recurrent Neural Network (RNN): The connection between nodes in an RNN results in a cycle such that the current input can affect the output from the next input, that is, the DNN has 'memory'. Accordingly, this type of NN is used for data that have a temporal correlation. In WCS, channel estimation has been addressed using RNN [3].

Overview on the Training of a Neural Network

The training of an NN, that is, the determination of the weights and biases of an NN, is not straightforward. In fact, the lack of an appropriate means of finding weights and biases resulted in a reduction in the volume of research on NNs after they were first developed in the 1950s. In the 1980s, an efficient method of training NN, that is, the back-propagation algorithm, was developed [2], leading to a second surge in research interest. The back-propagation algorithm is based on the gradient descent technique, such that the error of the NN, that is, the difference between the target and the output, is propagated in a backward direction to update the weights and biases according to the gradient, for example, to strengthen the parameters when the output is correct and to weaken them when the output is incorrect.

The training of an NN falls into three possible categories, as for the case of the training of general machine learning algorithms.

Supervised Learning (SL): In SL, the training data are labeled such that the NN can compare its output with ground-truth values. This type of learning is widely used in classification methods, for example, the detection of primary users in cognitive radio systems [9].

Unsupervised Learning (UL): In UL, the data are not labeled and the NN must autonomously derive meaningful features from input samples, for example, clustering. In [5] and [6], the encoder and decoder for optical wireless communications and SCMA systems were found using the autoencoder structure with UL.

Reinforcement Learning (RL): In RL, the learning of the NN is achieved by trial-and-error. In particular, for given input data, the proper action can be found by the NN, and a reward can be observed for the selected action. Then, the NN is trained to provide a better action, which leads to a higher reward. A multiple access based on RL was developed in [10].

Deep Neural Network and Wireless Communications Systems

For some time there was a belief that it was hard to train an NN with a large number of layers, even with the back-propagation algorithm, and

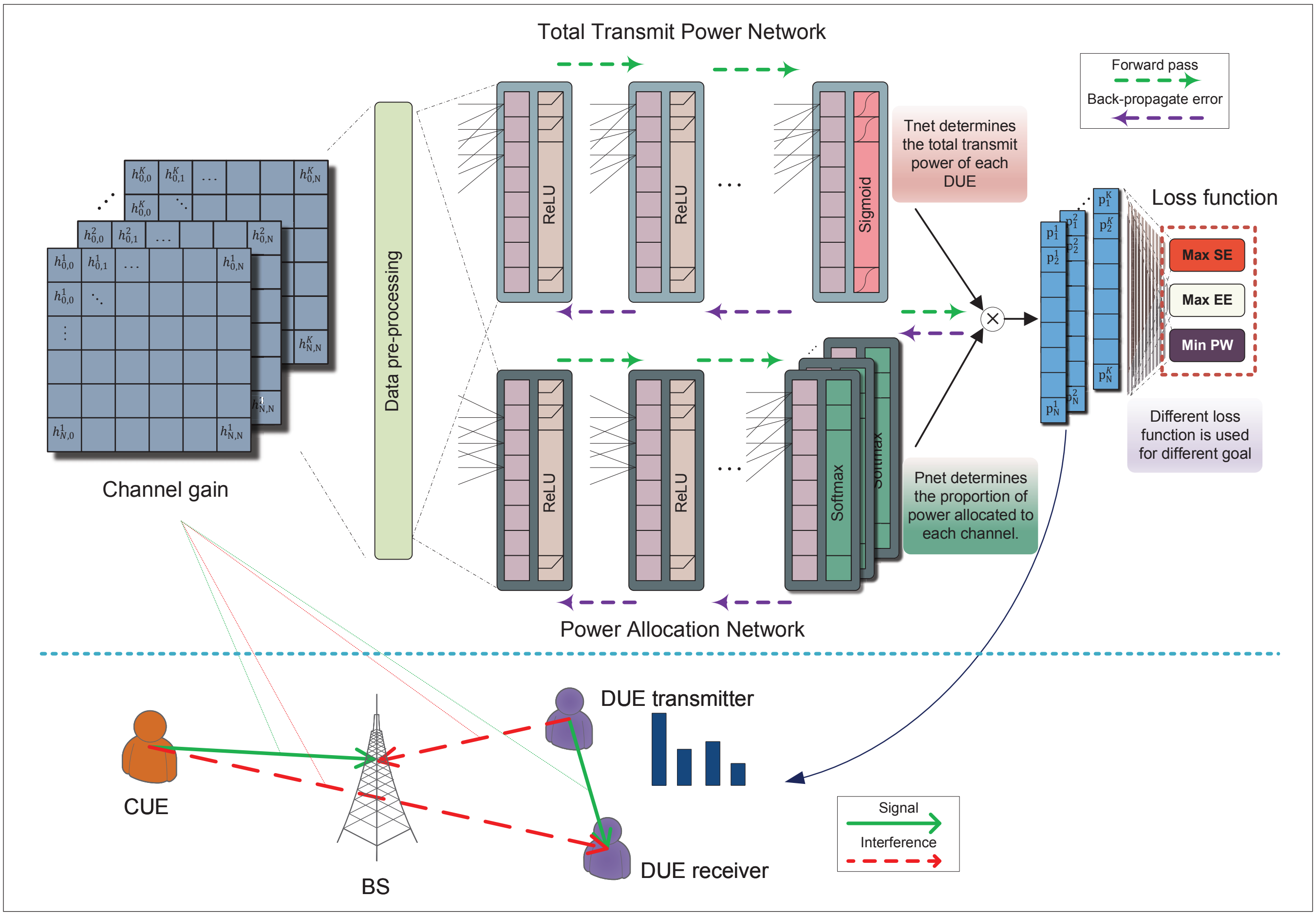


Figure 2. Considered system model and DNN structure for the RA.

this resulted in a second non-productive period in the research on NNs [2]. More recently, the use of NNs with a large number of layers, known as a DNN, has become possible, essentially due to the following four reasons [2].

Availability of a Large Dataset: Due to the development of sensors and the Internet, it is now possible to collect the large data sets required to allow a DNN to learn the general characteristics of a system. For instance, the development of ImageNet, which is a database containing more than 15,000,000 labeled images, has become the basis for the notable success of DNNs in image classification.

Use of Better Activation Functions: The use of piecewise linear activation functions, for example, rectified linear unit (ReLU) or leaky ReLU, in place of traditional activation functions such as sigmoid, has greatly improved the performance of DNNs [11, 12]. In particular, the use of piecewise linear activation functions can prevent the problem of the vanishing gradient of sigmoid, and is computationally more straightforward in that the DNN can be trained more efficiently.

Development of Parallel Computation: Although the inference of the output for one set of input data requires only a small computational overhead, the training can require a long computation time due to the large amount of training data used. However, thanks to the development of parallel computation based on graphics processing units, rather less computation time is now required.

Efficient Initialization Methodology: In DNN, the initialization technique is of the utmost importance in order to achieve high performance without the risk of encountering a local minimum hazard. Recently, effective initialization techniques such as Restricted Boltzmann Machine (RBM) and Xavier initialization enable the efficient training of DNN [12].

DNNs are suitable for WCS because they generally deal with a large amount of data and entail a complex system model that is hard to optimize analytically. In particular, unlike conventional approaches in WCS, where a specific channel model is assumed such that its performance cannot be guaranteed when the assumed channel differs from the actual channel, the DNN-based approach enables the adaptation of its operation according to the environment without relying on any specific channel model. Moreover, as for image classification, the DNN-based approach could yield better schemes than those based on conventional handcrafted techniques, especially when the problem is complex and hard to analyze [6, 8, 9].

Intelligent Resource Allocation Based on Deep Learning

In this section, we investigate intelligent RA using a DNN. To this end, we first describe our system model, which consists of multi-channel underlay device-to-device (D2D) communication. Although we consider a specific communication scenario, our system model comprises key considerations in general WCS, for example, inter-user interference, multi-channel usage, and a QoS requirement. We then discuss the proposed DNN model for RA, the objective of which is either to maximize SE or EE or to minimize the total transmit power, and the optimality and feasibility of DNN-based RA are then assessed via simulation.

System Model and Resource Allocation

We consider an RA for a multi-channel cellular system with underlay D2D communication, where data transmission of N D2D user equipments (DUEs) takes place at the same time to cellular transmission using K multi-channels. We assume that users are randomly distributed over an area $D \times D$. The channel gain between the i-th transmitter and the j-th receiver for channel k is denoted as $h^k_{i,j}$, where the index 0 is assigned to the cellular user equipment (CUE) and base station (BS). The system model is depicted in Fig. 2.

In the proposed RA, the transmit power of the DUEs allocated to each channel, which we denote as P^k_i where k and i are the index of the channel and the DUE, respectively, must be determined to maximize SE, maximize EE, or minimize the total transmit power. In the RA, we take three constraints into account. First, the transmit power allocated to each channel should be non-negative, and the sum of the transmit power for a single DUE must not exceed the maximum transmit power, P_T (i.e., the transmit power constraint). Second, the amount of interference caused to the cellular transmission should be less than threshold I_T (i.e., interference constraint). Third, each D2D transmission should satisfy the minimum QoS requirement, that is, the SE achieved by each DUE must be greater than threshold, R_T (i.e., QoS constraint). More detailed mathematical formulation can be found in [13].

Given that the considered RA is formulated into a non-convex problem, it is hard to derive the optimal RA analytically; therefore, iterative methods based on Lagrangian relaxation were considered [14]. However, this approach requires a number of iterative computations, which could increase the computation time [7, 8], such that real-time operation could be difficult. However, in DNN-based RA, the generic solver is derived autonomously by the DNN, which involves only simple matrix operations, such that the proper RA can be found for any channel condition with a low computational complexity without iteration.

Resource Allocation Based on Deep Neural Network

In our DNN model, the total transmit power of each DUE and the proportion of transmit power allocated to each channel by individual DUEs are found from the channel gain, $h^k_{i,j}$, such that the channel gain becomes the input of the DNN, and the transmit power and power allocation become the output of the DNN. Two separate DNN modules, namely the total transmit power network (Tnet) and the power allocation network (Pnet) are jointly used in our DNN model (see Fig. 2). The channel gain is pre-processed to improve its training performance such that it is converted to dB and normalized to ensure zero mean and unit variance [8]. These pre-processed channel gains become the input to the two networks, Tnet and Pnet.

Tnet and Pnet are composed of sub-modules, which consist of a fully connected (FC) layer and a ReLU used as an activation function. Given that each output of Tnet should be less than the maximum transmit power, P_T, we implemented a sigmoid, which is multiplied by the value of P_T at the end of Tnet, where the output of the sigmoid lies between 0 and 1. On the other hand, the proportion of transmit power allocated to each channel for each DUE is determined by Pnet, such that we consider N softmax modules as the last layer of Pnet, where each softmax module has K outputs. Given that the sum of a single softmax module is 1, it is appropriate to model the proportion of the transmit power allocated to channels for individual DUEs. Finally, by multiplying the outputs of Tnet and Pnet, the transmit power of DUEs allocated to each channel can be found. It should be noted that the same DNN structure with different values of weights and biases can be used for three different objectives, thereby making the practice efficient thanks to the reuse of the same DNN structure.

In order to find the optimal set of weights and biases, our proposed DNN model must first be trained. To this end, the channel samples for training must first be collected either by measurement or simulation using the well known channel model. Unlike for the SL based scheme [7], in our proposed scheme, the optimal transmit power for each channel sample is not needed for training due to the use of UL, which can greatly reduce the overhead in the preparation of the training data [8].

Then, the proposed DNN can be trained based on a back-propagation algorithm, where these channel samples become the input of the training. For the training, in order to achieve the goal while satisfying the constraints, the weighted sum of the objective function, that is, the overall SE, EE, or the total transmit power, and the functions regarding the constraints, are taken into account. In the loss function, the weight of the objective function should be negative for the maximization problem because the DNN model is trained to minimize the value of loss function. Moreover, the functions regarding the constraints should be the increasing function with respect to the violation of constraints, such that the value of the loss function for the DNN model decreases when the constraints are satisfied. Accordingly, the DNN model can be trained to achieve the desired goal, for example, the maximization of SE, while guaranteeing the constraints, that is, the interference constraint and the QoS constraint. The detailed formulation of loss functions can be found in [13].

After the proper loss function has been determined, the training of the DNN can be conducted efficiently using off-the-shelf stochastic gradient descent algorithms, for example, the Adaptive Moment Estimation algorithm. Subsequent to the training, the appropriate transmit power allocated to each channel can be determined by feeding the current channel gain to the trained DNN.

After the proper loss function has been determined, the training of the DNN can be conducted efficiently using off-the-shelf stochastic gradient descent algorithms, for example, the Adaptive Moment Estimation algorithm. Subsequent to the training, the appropriate transmit power allocated to each channel can be determined by feeding the current channel gain to the trained DNN.

Performance Evaluation

We now compare the performance of the proposed DNN-based RA with the optimal performance found using an exhaustive search (ES), in which the transmit power level of the DUE in

> We find that although DNN with SL has a simpler structure, the performance of our proposed scheme is better, and the overhead required to obtain the training data is much lower thanks to the use of UL. These simulation results affirm the applicability and benefits of our proposed DNN-based RA, resulting from its much lower computation time than the optimal scheme.

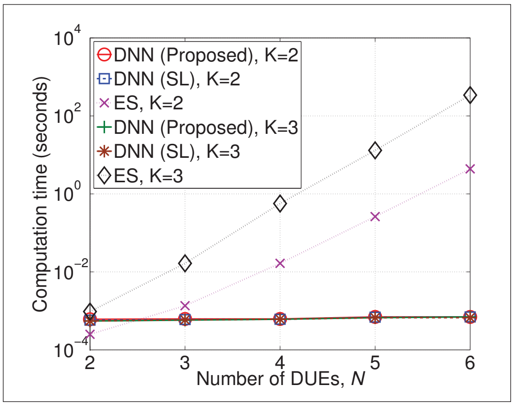


Figure 3. Computation time vs. number of DUEs.

using K multi-channels

each channel is quantized with Q equally spaced values, and all combinations of quantized values are examined. In the simulation, we also consider DNN-based RA using the SL proposed in [7] where the transmit power level of the ES is used as label data.

The maximum transmit power, the transmit power of the CUE, and the circuit power of the DUE are set to 20 dBm. Moreover, the bandwidth and the noise spectral density are set to 10 MHz and –174dBm/Hz, respectively. Furthermore, I_T = –50dBm, and R_T = 3 bps/Hz. In addition, the DUEs and CUEs are randomly distributed over an area in which the maximum distance between the transmitter and receiver of the same transmit pair is set at 15 m. We consider a simplified path loss model with path loss coefficient $10^{3.453}$ and path loss exponent 3.8, and an independent and identically distributed (i.i.d.) circularly symmetric complex Gaussian (CSCG) random variable is used for multipath fading, with zero mean and unit variance.

For the proposed DNN model, we assume that the number of layers in Tnet and Pnet is 4 and the number of hidden nodes in the FC layer is 100. Moreover, 40,000 channel samples are generated by the simulation and used for the training. Our simulation codes, together with further simulation results, can be found in [13].

First, we compare the computational complexity of the DNN-based RA schemes with that of the ES. It should be noted that the computational complexities of DNN-based RA and ES can be derived as $O(N^2K)$ and $O(Q^{NK})$, respectively, such that the computation time of ES increases more rapidly compared to that of DNN-based RA. In Fig. 3, we show the measured computation time for various values of N when K = 3 and Q = 5. As confirmed by the simulation results, the computation time of ES increases exponentially as K increases, while that of the DNN-based schemes remain at a low level, that is, the computation time of the DNN-based schemes is less than 0.7 milliseconds. Although we do not show this graphically, the computation times of our proposed scheme and DNN-based RA using SL for a large network configuration where $N = K$ = 50, are found to be 8.96 and 7.8 milliseconds, respectively, which are sufficiently low for real-time operation.

Next, we consider the average SE, the average EE, and the total transmit power of each DUE obtained for various values of area, D, as shown in Figs. 4a–c, respectively. Herein, we consider the proposed DNN-based RA, whose objectives are the maximization of SE (Max. SE), the maximization of EE (Max. EE), and the minimization of the total transmit power (Min. PW). Moreover, we also consider the ES with Q = 5, DNN with SL [7], and a scheme in which the transmit power is determined randomly, for comparative purposes. Furthermore, we assume that $N = K = 3$, such that finding the performance via ES is computationally plausible, in view of the fact that we were unable to obtain the performance of ES for a larger system configuration due to the excessive computation time. It is also worth noting that we were unable to check the performance of DNN with SL for a larger system configuration because the label data required for training must be obtained through ES.

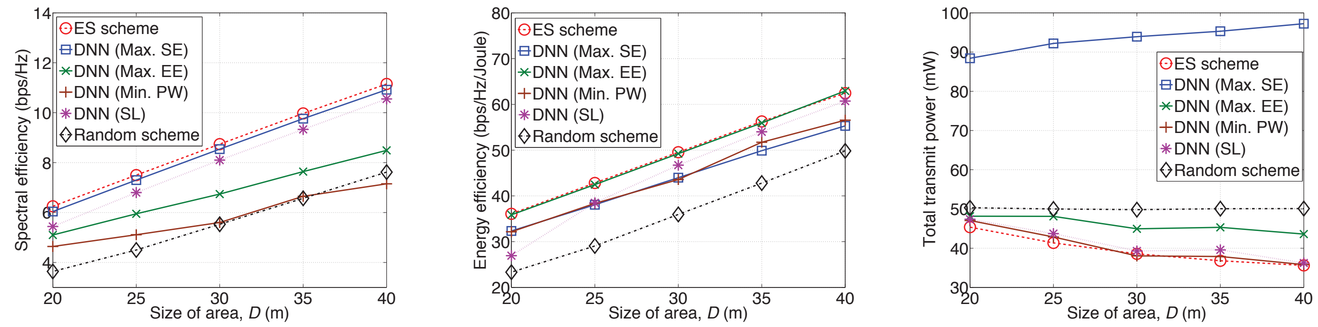
As shown by the simulation results, the random scheme shows the worst performance of all the considered schemes, which justifies the need for RA. We also find that our DNN-based RA produces close-to-optimal performance. For example, our proposed scheme achieves 96.6 percent of the SE of ES in all cases. Furthermore, it can be observed that SE and EE both increase with increasing area because the DUEs become more sparsely located, which results in a reduction of interference among DUEs. In addition, we find that although DNN with SL has a simpler structure, the performance of our proposed scheme is better, and the overhead required to obtain the training data is much lower thanks to the use of UL. These simulation results affirm the applicability and benefits of our proposed DNN-based RA, resulting from its much lower computation time than the optimal scheme.

Research Challenges

In the following, we discuss some of the research challenges connected with the future use of DL in WCS.

Measured Channel Data as Training Sample: The DNN finds the optimal strategy autonomously directly from the data, rather than using a handcrafted mathematical model. Accordingly, in order to achieve a high performance, a large number of training samples related to WCS, for example, channel gain data, must be collected for different scenarios using real measurements. The difficulties encountered in preparing the training data can be solved using DNN based on a generative model, for example, a generative adversarial network (GAN), to generate realistic synthetic training data.

Distributed Operation and Inaccurate Channel Information: Unlike in image classification where all input data can easily be obtained by a single node, in WCS it is hard to obtain the input data of the DNN, for example, the channel gain, by a single node executing the DNN functionality due to the high signaling overhead, especially when the number of users is large. Moreover, the input data of DNN for practical WCS is likely to contain errors due to inaccurate measurements and delayed channel feedback. To solve the problem of distributed operation, distributed DL can be considered [15], and to solve the problem of inaccurate channel information, a denoising autoencoder can be used.



(a) Spectral efficiency vs. size of area. (b) Energy efficiency vs. size of area. (c) Total transmit power vs. size of area.

Figure 4. Performance of DNN-based RA scheme.

Computation Complexity: The control of WCS, for example, RA, must be conducted within a very short time period, that is, a few milliseconds, due to the short frame length. Moreover, a scheme based on DL for WCS must have low computational complexity, allowing its use in a mobile device with limited computational power. Although we show that this DNN-based RA has a very low computation time compared with the optimal scheme, storing the trained DNN model can become an overhead for mobile devices. Furthermore, online learning, in which each mobile device trains its own model based on the channel samples collected, can be used to improve performance further, and in this case, the overhead associated with DNN-based schemes can become rather large. This problem can be solved using the newly developed AI accelerator chip, for example, the tensor processing unit (TPU), which is likely to be implemented in mobile devices of the future given the increased use of technologies based on DL.

Conclusions

In this article, we have discussed DL for WCS in preference to conventional handcrafted engineering based on mathematical modeling. In particular, we obtained an optimal strategy using intelligent RA based on DL, which is difficult to obtain via a conventional approach. It was confirmed by performance evaluation that our proposed DNN-based RA can achieve near-optimal performance with a low computation time for various criteria of RA. We have also outlined some of the challenges involved in the application of DL in research on WCS.

Acknowledgment

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(NRF-2018R1D1A1B07040796).

References

[1] A. Karpathy and L. Fei-Fei, "Deep Visual-Semantic Alignments for Generating Image Descriptions," *Proc. IEEE CVPR*, Boston, MA, USA, June 2015.
[2] Y. LeCun, Y. Bengio, and G. Hinton, "Deep Learning," *Nature*, vol. 521, no. 7553, May 2015, pp. 436–44.
[3] T. J. O'Shea, S. Hitefield, and J. Corgan, "End-to-End Radio Traffic Sequence Recognition with Recurrent Neural Networks," *Proc. IEEE GlobalSIP*, Washington, DC, USA, Dec. 2016.
[4] H. Ye, G. Y. Li, and B. H. Juang, "Power of Deep Learning for Channel Estimation and Signal Detection in OFDM Systems," *IEEE Wireless Commun. Lett.*, vol. 7, no. 1, Feb. 2018, pp. 114–17.
[5] H. Lee *et al.*, "Deep Learning Framework for Wireless Systems: Applications to Optical Wireless Communications," *IEEE Commun. Mag.*, vol. 57, no. 3, Mar. 2019, pp. 35–41.
[6] M. Kim *et al.*, "Deep Learning Aided SCMA," *IEEE Commun. Lett.*, vol. 22, no. 4, Apr. 2018, pp. 720–23.
[7] H. Sun *et al.*, "Learning to Optimize: Training Deep Neural Networks for Interference Management," *IEEE Trans. Signal Process.*, vol. 66, no. 20, Oct. 2018, pp. 5438–53.
[8] W. Lee, M. Kim, and D. H. Cho, "Deep Power Control: Transmit Power Control Scheme Based on Convolutional Neural Network," *IEEE Commun. Lett.*, vol. 22, no. 6, 2018, pp. 1276–79.
[9] W. Lee, M. Kim, and D. Cho, "Deep Cooperative Sensing: Cooperative Spectrum Sensing Based on Convolutional Neural Networks," *IEEE Trans. Veh. Technol.*, vol. 68, no. 3, Mar. 2019, pp. 3005–09.
[10] Y. Yu, T. Wang, and S. C. Liew, "Deep-Reinforcement Learning Multiple Access for Heterogeneous Wireless Networks," *IEEE JSAC*, vol. 37, no. 6, June 2019, pp. 1277–90.
[11] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, MIT Press, 2016.
[12] X. Glorot and Y. Bengio, "Understanding the Difficulty of Training Deep Feedforward Neural Networks," *Proc. AISTATS*, Sardinia, Italy, May 2010.
[13] "Simulation code," https://github.com/seotaijiya/IRAWCS; accessed: August 26, 2019.
[14] Y. Jiang *et al.*, "Energy-Efficient Joint Resource Allocation and Power Control for D2D Communications," *IEEE Trans. Veh. Technol.*, vol. 65, no. 8, Aug. 2016, pp. 6119–27.
[15] P. de Kerret and D. Gesbert, "Robust Decentralized Joint Precoding Using Team Deep Neural Network," *Proc. ISWCS*, Lisbon, Portugal, Aug. 2018.

Biographies

Woongsup Lee [S'07, M'13] (wslee@gnu.ac.kr) received a B.S. and a Ph.D. in electrical engineering from KAIST in 2006 and 2012, respectively. Since 2014, he has been working in the Department of Information and Communication Engineering of Gyeongsang National University, South Korea, where he is now an associate professor. His research interests include deep learning and future wireless communication systems.

Ohyun Jo [S'05, M'14] (ohyunjo@chungbuk.ac.kr) is currently an assistant professor in the Department of Computer Science at Chungbuk National University. He received his B.S., M.S., and Ph.D. degrees in electrical engineering from Korea Advanced Institute of Science and Technology (KAIST) in 2005, 2007, and 2011, respectively. His research interests include millimeter wave communications, next generation WLAN/WPAN systems, 5G mobile communication systems, military communications, Internet of Things, and communications FPGA/ASIC design.

Minhoe Kim [M'17] (kimminhoe@gmail.com) received his B.S., M.S., and Ph.D degrees in electrical engineering from Korea Advanced Institute of Science and Technology (KAIST) in 2012, 2014, and 2018 respectively. Since 2018, he has been with EURECOM, Sophia Antipolis, France, where he is a post-doctoral researcher. His main research interests include artificial intelligence, deep learning, beyond 5G wireless communication systems, and wireless energy harvesting.