

MACHINE LEARNING PARADIGMS FOR NEXT-GENERATION WIRELESS NETWORKS

CHUNXIAO JIANG, HAIJUN ZHANG, YONG REN, ZHU HAN,
KWANG-CHENG CHEN, AND LAJOS HANZO

ABSTRACT

Next-generation wireless networks are expected to support extremely high data rates and radically new applications, which require a new wireless radio technology paradigm. The challenge is that of assisting the radio in intelligent adaptive learning and decision making, so that the diverse requirements of next-generation wireless networks can be satisfied. Machine learning is one of the most promising artificial intelligence tools, conceived to support smart radio terminals. Future smart 5G mobile terminals are expected to autonomously access the most meritorious spectral bands with the aid of sophisticated spectral efficiency learning and inference, in order to control the transmission power, while relying on energy efficiency learning/inference and simultaneously adjusting the transmission protocols with the aid of quality of service learning/inference. Hence we briefly review the rudimentary concepts of machine learning and propose their employment in the compelling applications of 5G networks, including cognitive radios, massive MIMOs, femto/small cells, heterogeneous networks, smart grid, energy harvesting, device-to-device communications, and so on. Our goal is to assist the readers in refining the motivation, problem formulation, and methodology of powerful machine learning algorithms in the context of future networks in order to tap into hitherto unexplored applications and services.

INTRODUCTION

Radical and sometime even un-orthodox next-generation networking concepts have received substantial attention both in the academic as well as industrial communities. One of their driving forces is that of providing unprecedented data rates for supporting radical new applications. Specifically, next-generation networks are expected to learn the diverse and colorful characteristics of both the users' ambience as well as human behavior, in order to autonomously determine the optimal system configurations. These smart mobile terminals have to rely on sophisticated learning and decision-making. Machine learning, as one of the most powerful artificial intelligence tools, constitutes a promising solution [1]. As shown in Fig. 1, we may envision an intelligent radio that is capable of autonomously accessing the avail-

able spectrum with the aid of learning, altruistically controlling transmission power for the sake of conserving energy as well as adjusting the transmission protocols.

Machine learning has found wide-ranging applications in image/audio processing, finance and economics, social behavior analysis, project management, and so on [2]. Explicitly, a machine learns the execution of a particular task \mathbb{T} , with the goal of maintaining a specific performance metric \mathbb{P} , based on a particular experience \mathbb{E} , where the system aims to reliably improve its performance \mathbb{P} while executing task \mathbb{T} , again by exploiting its experience \mathbb{E} . Depending on how we specify \mathbb{T} , \mathbb{P} , and \mathbb{E} , the learning might also be referred to as data mining, autonomous discovery, database updating, programming by example, and so on [3]. Machine learning algorithms can be simply categorized as supervised and unsupervised learning, where the adjectives "supervised/unsupervised" indicate whether there are labeled samples in the database. Later, reinforcement learning emerged as a new category that was inspired by behavioral psychology. It is concerned with an agent's certain form of reward/utility, who is connected to its environment via perception and action. The family of machine learning algorithms can also be categorized based on their similarity in terms of their functionality and structure, yielding regression algorithms, instance-based algorithms, regularization algorithms, decision tree algorithms, Bayesian algorithms, clustering algorithms, association rule based learning algorithms, artificial neural networks, deep learning algorithms, dimension reduction algorithms, ensemble algorithms, and so on. In this article, we will introduce the basic concept of machine learning algorithms and the corresponding applications according to the category of supervised, unsupervised, and reinforcement learning.

Machine learning can be widely used in modeling various technical problems of next-generation systems, such as large-scale MIMOs, device-to-device (D2D) networks, heterogeneous networks constituted by femtocells and small cells, and so on. Figure 2 portrays the family-tree of machine learning techniques and their potential applications in 5G. Against this background, we embark on investigating the family of learning techniques. Specifically, in the following sections we consider supervised learning, unsupervised learning, and

Chunxiao Jiang is with the Tsinghua Space Center.

Y. Ren is with Tsinghua University.

Haijun Zhang is with the University of Science and Technology Beijing, China

Zhu Han is with the University of Houston.

Kwang-Cheng Chen is with the University of South Florida

Lajos Hanzo is with the University of Southampton.

Digital Object Identifier: 10.1109/MWC.2016.1500356WC

reinforcement learning. Each section consists of several subsections, discussing specific learning models, such as regression models and the k -nearest neighbor (KNN) algorithm, support vector machines (SVM) and Bayesian learning; k -means clustering, principal and independent component analysis; and partially observed Markov decision processes, Q-learning, and the multi-armed bandit technique. Each section commences with the introduction of the learning model and its applications in 5G networks. Finally, our conclusions are drawn.

SUPERVISED LEARNING IN WIRELESS COMMUNICATIONS

REGRESSION MODELS, KNN AND SVM: MIMO CHANNEL AND ENERGY LEARNING

Models: Regression analysis relies on a statistical process for estimating the relationships among variables. The goal of regression analysis is to predict the value of one or more continuous-valued estimation targets, given the value of a D -dimensional vector \mathbf{x} of input variables. The estimation target is a function of the independent variables. In linear regression, the regression function is linear, while in logistic regression, it is a logistic function assuming a common sigmoid curve. The KNN and SVM algorithms are mainly utilized for classification of points/objects. In KNN, an object is classified into a specific category by a majority vote of the object's neighbors, with the object being assigned to the class that is most common among its k nearest neighbors. The output may be constituted by a specific property of the object, such as for example the average of the values of its k nearest neighbors. By contrast, the SVM algorithm relies on nonlinear mapping, which transforms the original training data into a higher dimension where it becomes separable and then it searches for the optimal linear separating hyperplane that is capable of separating one class from another, again in this higher dimension. They correspond to non-linear classification methods relying on the family of kernel methods. It was

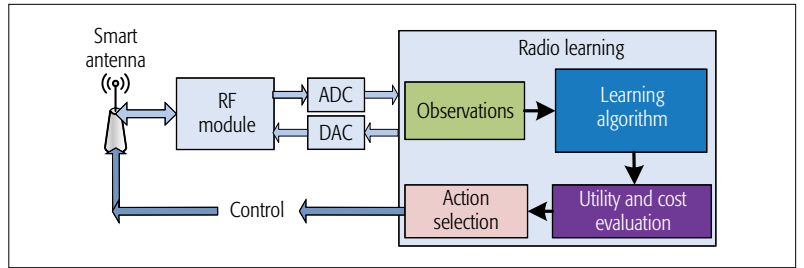


FIGURE 1. Intelligent radio learning paradigm.

shown that with the aid of an appropriate nonlinear mapping to a sufficiently high dimension, the data from two classes can always be separated by a hyperplane [3 p. 21, 185, 239, 349].

Applications: These models can be used for estimating or predicting radio parameters that are associated with specific users. For example, in massive MIMO systems associated with hundreds of antennas, both detection and channel estimation lead to high-dimensional search-problems, which can be addressed by the above-mentioned learning models. In order to generalize the SVM function for employment in data classification problems, its hierarchical version, referred to as H-SVM, was proposed in [4], where each hierarchical level consisted of a finite number of SVM classifiers. This regime was used for the estimation of the Gaussian channel's noise level in a MIMO-aided wireless network having t transmit antennas and r receive antennas. By exploiting the training data, the H-SVM model was trained for the estimation of the channel noise statistics.

In heterogeneous networks constituted by diverse cells, handovers may be frequent, where both the KNN and SVM can be applied to finding the optimal handover solutions. At the application layer, these models can also be used for learning the mobile terminal's specific usage pattern in diverse spatio-temporal and device contexts, as discussed in [5]. This may then be exploited for prediction of the configuration to be used in the location-specific interface. Given a set of contex-

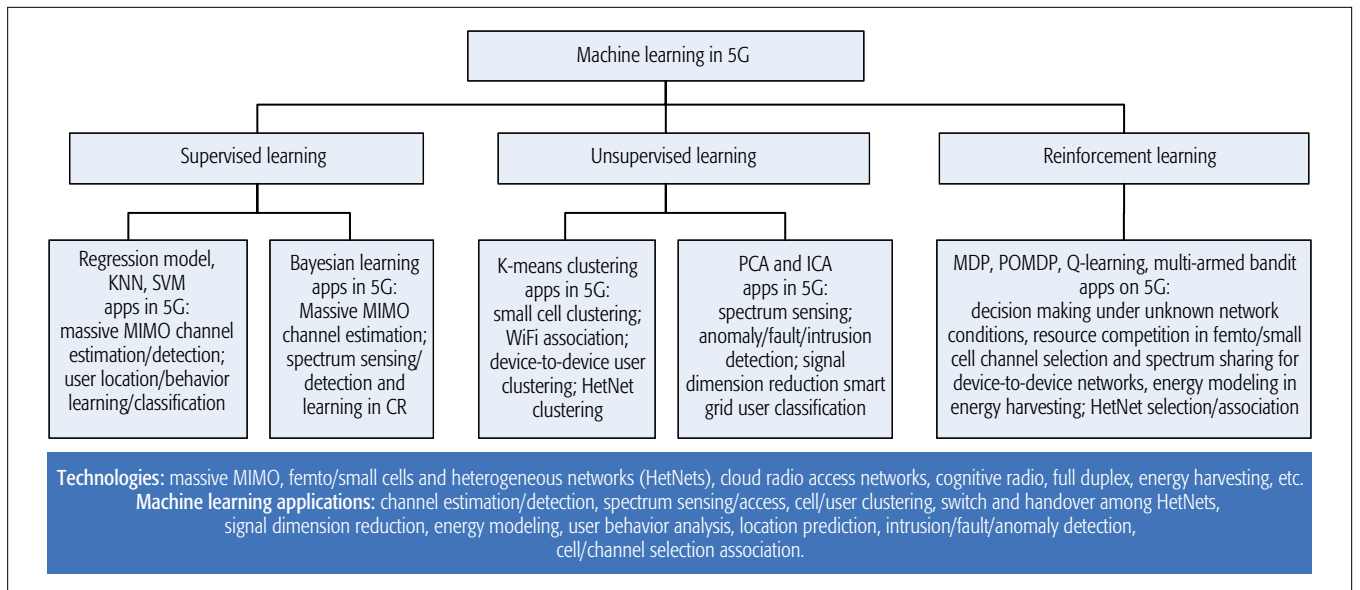


FIGURE 2. Radio learning architecture.

HMM is a tool designed for representing probability distributions of sequences of observations. It can be considered a generalization of a mixture-based model, where the hidden variables, which control the specific mixture of the component to be selected for each observation, are related to each other through a Markov process, rather than being independent of each other.

Category	Learning techniques	Key characteristics	Application in 5G
Supervised learning	Regression models	<ul style="list-style-type: none"> • Estimate the variables' relationships • Linear and logistics regression 	Energy learning [5]
	K-nearest neighbor	<ul style="list-style-type: none"> • Majority vote of neighbors 	Energy learning [5]
	Support vector machines	<ul style="list-style-type: none"> • Non-linear mapping to high dimension • Separate hyperplane classification 	MIMO channel learning [4]
	Bayesian learning	<ul style="list-style-type: none"> • <i>A posteriori</i> distribution calculation • GM, EM, and HMM 	<ul style="list-style-type: none"> • Massive MIMO learning [6] • Cognitive spectrum learning [7–9]

TABLE 1. Supervised machine learning algorithms.

tual input cues, machine learning algorithms are capable of exploiting the user context learned for the sake of dynamically classifying the cues into a system state for the sake of saving energy, while maintaining a high level of user satisfaction. Donohoo *et al.* [5] also conducted experiments using five real user profiles, including the user-locations and energy consumption, but their data is not accessible to the public. The experiment showed that up to 90 percent successful energy demand prediction is possible with the aid of the KNN algorithms.

BAYESIAN LEARNING: MASSIVE MIMO AND COGNITIVE RADIO

Models: The philosophy of Bayesian learning is to compute the a posteriori probability distribution of the target variables conditioned on its input signals and on all of the training instances. Some simple examples of generative models that may be learned with the aid of Bayesian techniques include, but are not limited to, the Gaussians mixture model (GM), expectation maximization (EM), and hidden Markov models (HMM) [3 p. 445].

GM is a model where each data point belongs to one of several clusters or groups, and the data points within each cluster are Gaussian distributed.

EM is a generalization of maximum likelihood estimation, which iteratively finds the most likely solutions or parameters. It is characterized by two steps: the “E” step that chooses a function representing the lower bound of the likelihood, and the “M” step that finds the parameters maximizing the chosen function.

HMM is a tool designed for representing probability distributions of sequences of observations. It can be considered a generalization of a mixture-based model, where the hidden variables, which control the specific mixture of the component to be selected for each observation, are related to each other through a Markov process, rather than being independent of each other.

Applications: The Bayesian learning model may be readily invoked for spectral characteristic learning and estimation in next-generation networks. To address the pilot contamination problem encountered in massive MIMO systems, the authors of [6] estimated both the channel parameters of the desired links in a target cell as well as those of the interfering links of the adjacent cells, where channel estimation was carried out with the aid of sparse Bayesian learning techniques. Based on the observation of received signals, the channel component was first modeled by a GM,

namely by a weighted sum of Gaussian distributions having different variances, and then estimated with the aid of the EM algorithm.

Another three closely related applications may be found in cognitive radio networks. In [7], a cooperative wideband spectrum sensing scheme based on the EM algorithm was proposed for the detection of a primary user (PU) supported by a multi-antenna assisted cognitive radio network. This iterative technique first created the log-likelihood function of both the unknown spectrum occupancy as well as of the channel information and of the noise in the “E” step. Then, it maximized the log-likelihood function for the sake of inferring the unknown information during the “M” step, which was carried out by jointly detecting both the PU signal as well as estimating the channel’s unknown frequency response and the noise variance of multiple subbands.

In contrast to [7], the authors in [8] constructed a HMM relying on a two-state hidden Markov process, where the PUs are present or absent and a two-state observation space, indicating whether the PUs are present or absent. Furthermore, the EM algorithm was invoked for finding the true channel parameters, such as the sojourn time of the available channels, the inactive states of the PUs, and the PUs’ signal strength. Finally, the third application of Bayesian learning was advocated in [9], where a tomography model, belonging to the Bayesian inference framework, was proposed for conceiving and statistically characterizing a range of techniques that are capable of extracting the prevalent parameters and traffic/interference patterns for employment in cognitive radio networks at both the link layer and network layer. The parameters collected included both the path-delay as well as the proportion of successful packet receptions, while the estimated parameter was the link’s successful transmission probability. The Bayesian estimators were derived for single/multiple transmissions in single/multiple path scenarios. In Table 1, we summarize the basic characteristics and applications of supervised machine learning algorithms.

UNSUPERVISED LEARNING IN WIRELESS COMMUNICATIONS

K-MEANS CLUSTERING: HETEROGENEOUS NETWORKS

Models: K-means clustering aims for partitioning n observations into k clusters, where each observation belongs to the closest cluster. It defines the centroid of a cluster as the center of

gravity, that is, the mean value of the points within the cluster. The clustering algorithm proceeds in an iterative manner, where an object is assigned to the specific cluster whose centroid is nearest to the object based on the Euclidean distance ‘similarity metric’, and then the in-cluster differences are minimized by iteratively updating the cluster-centroid, until ‘convergence’ is achieved. Explicitly, convergence is deemed to be achieved when the assignment becomes stable, that is, the clusters formed in the current round are the same as those formed in the previous round [3 p. 161, 317].

Applications: Clustering is a common problem in 5G networks, especially in heterogeneous scenarios associated with diverse cell sizes as well as WiFi and D2D networks. For example, the small cells have to be carefully clustered to avoid interference using coordinated multi-point transmission (CoMP), while the mobile users are clustered to obey an optimal offloading policy, the devices are clustered in D2D networks to achieve high energy efficiency, the WiFi users are clustered to maintain an optimal access point association, and so on. In [10], the authors considered a hybrid optical/wireless network scenario, in order to reduce the overall wireless tele-traffic by encouraging the utilization of the high-capacity optical infrastructure. A mixed integer programming (MIP) problem was formulated to jointly optimize both the gateway partitioning and the virtual-channel allocation based on classic k -means clustering, which was employed to partition the mesh access points (MAPs) into several groups. The proposed scheme commenced its operation from an initial gateway access point (GAP) set, which can be plucked by a random selection from the set of MAPs, or can be more astutely determined using a meritorious initialization criterion. Next, each MAP is assigned to its nearest GAP. If several eligible GAPs are in the vicinity, then the specific GAP that has a readily available virtual channel is chosen. Finally, by using the classic k -means clustering algorithm, the MAPs are divided into k groups associated with the closest GAPs.

PRINCIPAL AND INDEPENDENT COMPONENT ANALYSIS: SMART GRID AND COGNITIVE RADIO

Models: Principal component analysis (PCA) transforms a set of potentially correlated variables into a set of uncorrelated variables, referred to as the principal components, where the number of principal components is less than or equal to the number of original variables. Basically, the first principal component has the largest possible variance (i.e., accounts for as much of the variability in the data as possible), and each succeeding component in turn has the highest variance possible under the constraint that it is orthogonal to (i.e., uncorrelated with) the preceding components. The principal components are orthogonal, because they are the eigenvectors of the covariance matrix, which is symmetric. By contrast, independent component analysis (ICA) is a statistical technique conceived to reveal hidden factors that underlie sets of random variables, measurements, or signals. In the model, the data variables are assumed to be linear mixtures of some unknown latent variables, and the mixing system is also unknown. The latent variables are assumed to be

non-Gaussian and mutually independent, and they are referred to as the independent components of the observed data, which can be found by ICA [3 p. 115].

Applications: Both the PCA and ICA constitute powerful statistical signal processing techniques devised to recover statistically independent source signals from their linear mixtures. One of their major applications may be found in the area of anomaly-detection, fault-detection, and intrusion-detection problems of wireless networks, which rely on traffic monitoring. Furthermore, similar problems may also be solved in sensor networks, mesh networks, and so on. They can also be invoked for the physical layer signal dimension reduction of massive MIMO systems or to classify the primary users’ behaviors in cognitive radio networks. As a further example, in [11] PCA and ICA were applied in a smart grid scenario to recover the simultaneous wireless transmissions of smart utility meters installed in each home. At the power utility station, it was required to separate the signals received from all the smart meters before the signals can be decoded. The statistical properties of the signals were exploited to blindly separate them using ICA. This operation is capable of enhancing both the transmission efficiency by avoiding channel estimation in each frame, as well as data security by eliminating any wideband interference or jamming signals. More explicitly, a substantial security enhancement was achieved by a robust version of the PCA-based method, which exploited the sparse, low-rank nature of the auto-covariance matrices of the smart metering signal and of the wideband interferer, respectively, in order to confidently separate them prior to ICA processing. Another pertinent example is found in cognitive radio scenarios, where the so called Boolean ICA relied on the Boolean mixing of OR, XOR, and other functions of binary signals [12]. It was also incorporated into the PU separation problem often encountered in cognitive radio networks for the sake of distinguishing and characterizing the activities of PUs in the context of collaborative spectrum sensing. Furthermore, the observations of the secondary users (SUs) were modeled as Boolean OR mixtures of the underlying binary PU sources. An iterative algorithm, called Binary ICA, was developed to determine the activities of the underlying latent signal sources, such as the PUs. It was demonstrated that given m monitors or SUs, the activities of up to $(2m - 1)$ distinct PUs can be inferred. In Table 2, we summarize the basic characteristics and applications of unsupervised machine learning algorithms.

REINFORCEMENT LEARNING IN WIRELESS COMMUNICATIONS

PARTIALLY OBSERVABLE MARKOV DECISION PROCESS: ENERGY HARVESTING

Models: Markov decision processes (MDPs) provide a mathematical framework for modeling decision making in specific situations, where the outcomes are partly random and partly under the control of a decision maker, as illustrated in Fig. 3a. At each time step, the process is in some state s , and the decision maker may opt for any

Principal component analysis (PCA) transforms a set of potentially correlated variables into a set of uncorrelated variables referred to as the principal components, where the number of principal components is less than or equal to the number of original variables.

Category	Learning techniques	Key characteristics	Application in 5G
Unsupervised learning	K-means clustering	<ul style="list-style-type: none"> • K partition clustering • Iterative updating algorithm 	Heterogeneous networks [10]
	PCA	<ul style="list-style-type: none"> • Orthogonal transformation 	Smart grid [11]
	ICA	<ul style="list-style-type: none"> • Reveal hidden independent factors 	Spectrum learning in cognitive radio [12]

TABLE 2. Unsupervised machine learning algorithms.

of the legitimate actions a that is available in state s . The process responds at the next time step by randomly moving into a new state s' , and giving the decision maker a corresponding reward $U_a(s)$. The probability that the process moves into its new state s' is influenced both by the specific action chosen, as well as by the system's inherent transitions, formally described by the state transition probability $P_a(s'|s, a)$. Given s and a , the state transition probability is conditionally independent of all previous states and actions, that is, the state transitions of an MDP process satisfy the fundamental Markov property. By contrast, a partially observable Markov decision process (POMDP) may be viewed as the generalization of a MDP, where the agent is unable to directly observe the underlying state transitions and hence only has partial knowledge, as shown in Fig. 3b. The agent has to keep track of both the probability distribution of the legitimate states, based on a set of observations, as well as of the observation probabilities and of the underlying MDP [3 p. 517].

Applications: The family of MDP/POMDP models constitutes ideal tools for supporting decision making in 5G networks, where the users may be regarded as agents and the network constitutes the environment. There are usually three steps associated with modeling a problem using MDP. The first step is to specify the system's state space and the decision maker's action space, as well as verifying the Markov property. The second step is that of constructing the state transition probabilities $P_a(s'|s, a)$ formulated as the probability of traversing from state s to s' under action a . The last step is to quantify both the decision maker's immediate reward $U_a(s)$ and its long-term reward using Bellman's equation [13]. Then, a carefully constructed iterative algorithm may be conceived to identify the optimal action in each state.

Classical applications found in the literature include the network selection/association problems of heterogeneous networks (HetNets), channel sensing, and user access in cognitive radio networks, and so on. Furthermore, energy harvesting (EH) has also been extensively modeled using MDP/POMDP, where the limited battery and the time-variant channels are usually regarded as the environment, while the users' channel selection or battery utilization are usually considered as the actions. For instance, in [13] the transmission power control problems of EH systems were investigated using the POMDP model, where the state space was defined by including the battery state, the channel state, the packet transmission/reception states, and an action by the node, which corresponded to sending a packet at a certain power level. The feedback messages implicitly provided the EH system with partial channel state information (CSI), which resulted in the corresponding POMDP formulation. Since finding exact solutions to the POMDP tends to be computationally intractable [13], a pair of computationally efficient suboptimal solutions, i.e. the maximum-likelihood heuristic policy and the voting heuristic policy, were explored.

Q-LEARNING: FEMTO/SMALL CELLS

Models: Q-learning may be invoked to find an optimal action policy for any given (finite) Markov decision process, especially when the system model is unknown, as shown in Fig. 3c. It is a model-free reinforcement learning technique and as such it can be used in conjunction with MDP models. In such a case, the Q-learning model is also comprised of an agent, of the states S and of a set of actions A per state. By executing an action in a specific state, the agent gleans a reward and the goal is to maximize its accumulated reward. Such a reward is illustrated by a Q-function, where "Q" is initialized to be an (arbitrary) fixed value. Then, "Q" is updated in an iterative manner after the agent carries out an action and observes the resultant reward as well as the associated new state at each time-instant [3 p. 517].

Applications: Q-learning has also been extensively applied in heterogeneous networks, usually in conjunction with the aforementioned MDP models. In [14] the authors presented a heterogeneous fully distributed multi-objective strategy based on a reinforcement learning model con-

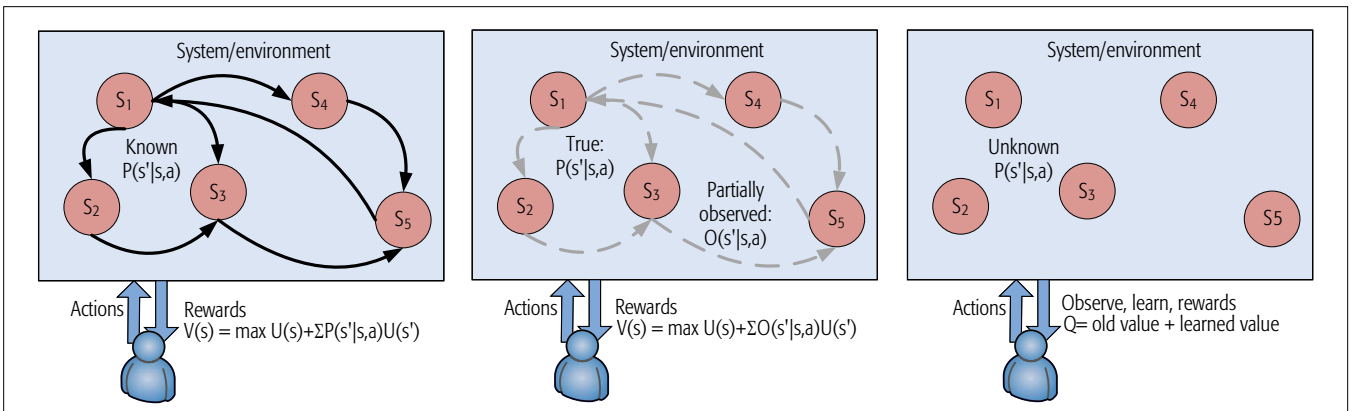


FIGURE 3. Illustration of reinforcement learning: a) Markov decision process; b) partially observed Markov decision process; c) Q-learning.

structured for the self-configuration/optimization of femtocells. The model was supposed to solve both the resource allocation and interference coordination problems in the downlink of femtocell networks. The main objectives of the learning process are two-fold: first, to acquire spectrum allocation awareness and to identify the availability of unused spectral slots for the provision of opportunistic access; second, to select sub-channels from the available spectrum pool and to configure the terminals supported by femtocells to operate under carefully constructed restrictions to avoid interference and to meet the quality of service (QoS) requirements. Another example is constituted by dense small cell networks regarding their cell outage management and compensation [15]. The system's state was constituted by the specific allocation of users to the resource blocks of small cells, as well as by the channel quality, while the actions were constituted by the downlink power control actions, with the rewards being quantified in terms of signal-to-interference-plus-noise ratio (SINR) improvement. It was demonstrated that the compensation strategy based on the reinforcement learning model attained an exceptional performance improvement.

MULTI-ARMED BANDITS: DEVICE-TO-DEVICE NETWORKS

Models: In practice, multi-armed bandits (MAB) have been used to model resource allocation problems operating under a fixed budget by carefully proportioning resources among competing projects, whose properties are only partially known at the time of resource allocation, but which may become better understood as time passes. Since the agent has no initial knowledge about the machines, the crucial trade-off they confront at each instance is between the "exploitation" of the specific machine that has the highest expected payoff and the "exploration" required to glean more information about the expected payoffs of the other machines.

The MAB problem may also be extended into a multi-player, multi-armed bandit game (MP-MAB), where the reward gleaned by any player depends on the specific decisions of other players. The key idea of the proposed approach is to enable each user to forecast the future actions of its opponents based on public knowledge and to proceed by best responding to the predicted joint action profile using some bandit strategy [3 p. 517].

Applications: The MAB and MP-MAB models, as a family of emerging signal processing tools, are capable of solving challenging resource allocation problems in wireless scenarios, where either the channel conditions or some other wireless environment parameters have to be "explored," while the known channels also have to be "exploited" by a group of users. Generally, these models may be beneficially used in multi-player adaptive decision making problems, where selfish players infer an optimal joint action profile from their successive interactions with a dynamic environment, and finally settle at some equilibrium point. This problem has indeed been encountered in many wireless networking scenarios, with a compelling one being the channel selection problem of a distributed device-to-de-

Category	Learning techniques	Key characteristics	Application in 5G
Reinforcement learning	MDP/POMDP	<ul style="list-style-type: none"> • Bellman equation maximization • Value iteration algorithm 	Energy harvesting [13]
	Q-learning	<ul style="list-style-type: none"> • Unknown system transition model • Q-function maximization 	Femto and small cells [14, 15]
	Multi-armed bandit	<ul style="list-style-type: none"> • Exploration vs. exploitation • Multi-armed bandit game 	Device-to-device networks [16]

TABLE 3. Reinforcement machine learning algorithms.

vice (D2D) communication system integrated into a cellular network, and another one in the context of emerging next-generation networks [16]. The selfish D2D users aimed to optimize their own performance by camping on the vacant cellular channels, whose statistics were unknown to the users. This distributed channel selection problem was in harmony with the typical MP-MAB settings, and thus it was modeled as an MP-MAB game. Specifically, every D2D user was modeled as a player of the MP-MAB game, while the channels were regarded as arms and choosing a channel corresponds to pulling an arm. The authors proposed a channel selection strategy consisting of two main blocks, namely the calibrated forecasting and the no-regret bandit learning strategies. In Table 3, we summarize the rudimentary characteristics and applications of reinforcement machine learning algorithms.

FUTURE RESEARCH AND CONCLUSIONS

A range of future research ideas on machine learning in 5G networks can be summarized as follows.

The family of supervised learning techniques relies on known models and labels that can support the estimation of unknown parameters. They can be utilized for massive MIMO channel estimation and data detection, spectrum sensing and white space detection in cognitive radio, as well as for adaptive filtering in signal processing for 5G communications. They can also be applied in higher-layer applications, such as inferring the mobile users' locations and behaviors, which can assist the network operators to improve the quality of their services.

Unsupervised learning relies on the input data itself in a heuristic manner. It can be utilized for cell clustering in cooperative ultra-dense small-cell networks, for access point association in ubiquitous WiFi networks, for heterogeneous base station clustering in HetNets, and for load-balancing in HetNets. It can also be applied in anomaly/fault/intrusion detection and for the users' behavior-classification.

Reinforcement learning relies on a dynamic iterative learning and decision-making process. It can be utilized for inferring the mobile users' decision making under unknown network conditions, for example during channel access under unknown channel availability conditions in spectrum sharing, for distributed resource allocation under unknown resource quality conditions in femto/small-cell networks, and base station asso-

The classes of supervised, unsupervised, and reinforcement learning tools were investigated, along with the corresponding modeling methodology and possible future applications in 5G networks. In a nutshell, machine learning is an exiting area for artificial intelligence aided networking research!

ciation under the unknown energy status of the base stations in energy harvesting networks.

Furthermore, computational intelligence paradigms, such as neural networks and neuro-fuzzy methods, swarm intelligence algorithms such as ant colony optimization, and evolutionary algorithms such as the competitive imperialist algorithm, may also be applied to improve the performance of 5G networks. Among those compelling techniques, neural networks and deep learning have recently become particularly popular. Generally, a neural network consists of a number of neurons and weighted connections among them, where the neurons can be regarded as variables and the weights can be viewed as parameters. The network should be appropriately configured with the aid of learning techniques to ensure that the application of a set of inputs produces the desired set of outputs. Explicitly, this can be achieved by iteratively adjusting the weights of the existing connections among all neuron pairs with the aid of learning based on the labeled data for supervised learning or unlabeled data for unsupervised learning. Neural networks have been widely utilized for spectral white state estimation [17], prediction [18], and handoff decisions [19] in cognitive radio networks. Note that the algorithms introduced in this article are only limited samples of the machine learning field. There are many other algorithms that can also be applied to the next-generation networks. For example, the family of evolutionary algorithms, such as genetic algorithms can solve optimization problems by mimicking a natural selection process, which can be utilized to solve resource allocation problems in HetNets [20]. By contrast, machine learning relies on two phases, the training phase and the testing phase, where the training phase imposes a much higher complexity than the testing phase. Due to the energy constraints and computational complexity constraints of mobil terminals, it is recommended to only implement the testing phase on shirt-pocket-sized mobile terminals.

This article reviewed the benefits of artificial intelligence aided wireless systems equipped with machine learning. We introduced the major families of machine learning algorithms and discussed their applications in the context of next-generation networks, including massive MIMOs, the smart grid, cognitive radios, heterogeneous networks, femto/small cells, D2D networks, and so on. The classes of supervised, unsupervised, and reinforcement learning tools were investigated, along with the corresponding modeling methodology and possible future applications in 5G networks. In a nutshell, machine learning is an exiting area for artificial intelligence aided networking research!

ACKNOWLEDGMENT

This research was supported by NSFC China under projects 61371079, 61471025, and 91338203, by the Open Research Fund of the National Mobile Communications Research Laboratory, Southeast University (No. 2016D07), and also by a Postdoctoral Science Foundation funded project.

REFERENCES

[1] M. van der Schaar and F. Fu, "Spectrum Access Games and Strategic Learning in Cognitive Radio Networks for Delay-Critical Applications," *Proc. IEEE*, vol. 97, no. 4, Apr. 2009, pp. 720–40.

[2] M. J. Er and Y. Zhou, "Theory and Novel Applications of Machine Learning," *InTech*, 2009.

[3] E. Alpaydm, *Introduction to Machine Learning*, 3rd ed., The MIT Press, Cambridge, Massachusetts, 2014.

[4] P. Zhou, Y. Chang, and J. A. Copeland, "Determination of Wireless Networks Parameters through Parallel Hierarchical Support Vector Machines," *IEEE Trans. Parallel Distrib. Syst.*, vol. 23, no. 3, Mar. 2012, pp. 505–12.

[5] B. K. Donohoo et al., "Context-Aware Energy Enhancements for Smart Mobile Devices," *IEEE Trans. Mobile Comput.*, vol. 13, no. 8, Aug. 2014, pp. 1720–32.

[6] C.-K. Wen et al., "Channel Estimation for Massive MIMO Using Gaussian-Mixture Bayesian Learning," *IEEE Trans. Wireless Commun.*, vol. 14, no. 3, Mar. 2015, pp. 1356–68.

[7] K. W. Choi and E. Hossain, "Estimation of Primary User Parameters in Cognitive Radio Systems via Hidden Markov Model," *IEEE Trans. Signal Process.*, vol. 61, no. 3, Feb. 2013, pp. 782–95.

[8] A. Assra, J. Yang, and B. Champagne, "An EM Approach for Cooperative Spectrum Sensing in Multi-Antenna CR Networks," to appear in *IEEE Trans. Veh. Technol.*; DOI: 10.1109/TVT.2015.2408369, 2015.

[9] C.-K. Yu, K.-C. Chen, and S.-M. Cheng, "Cognitive Radio Network Tomography," *IEEE Trans. Veh. Technol.*, vol. 59, no. 4, May 2010, pp. 1980–97.

[10] M. Xia et al., "Optical and Wireless Hybrid Access Networks: Design and Optimization," *OSA/IEEE J. Opt. Commun. Netw.*, vol. 4, no. 10, Oct. 2012, pp. 749–59.

[11] R. C. Qiu et al., "Cognitive Radio Network for the Smart Grid: Experimental System Architecture, Control Algorithms, Security, and Microgrid Testbed," *IEEE Trans. Smart Grid*, vol. 2, no. 4, Dec. 2011, pp. 724–40.

[12] H. Nguyen et al., "Binary Inference for Primary User Separation in Cognitive Radio Networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 4, Apr. 2013, pp. 1532–42.

[13] A. Aprem, C. R. Murthy, and N. B. Mehta, "Transmit Power Control Policies for Energy Harvesting Sensors with Retransmissions," *IEEE J. Sel. Topics Signal Process.*, vol. 7, no. 5, Oct. 2013, pp. 895–906.

[14] G. Alnwaimi, S. Vahid, and K. Moessner, "Dynamic Heterogeneous Learning Games for Opportunistic Access in LTE-Based Macro/Femtocell Deployments," *IEEE Trans. Wireless Commun.*, vol. 14, no. 4, Apr. 2015, pp. 2294–2308.

[15] O. Onireti et al., "A Cell Outage Management Framework for Dense Heterogeneous Networks," *IEEE Trans. Veh. Technol.*, vol. 65, no. 4, 2016, pp. 2097–2113; DOI: 10.1109/TVT.2015.2431371.

[16] S. Maghsudi and S. Stanczak, "Channel Selection for Network-Assisted D2D Communication via No-Regret Bandit Learning with Calibrated Forecasting," *IEEE Trans. Wireless Commun.*, vol. 14, no. 3, Mar. 2015, pp. 1309–22.

[17] K. Tsagkaris, A. Katidiotis, and P. Demestichas, "Neural Network-Based Learning Schemes for Cognitive Radio Systems," *Computer Commun.*, vol. 31, no. 14, Sep. 2008, pp. 3394–3404.

[18] V. K. Tumuluru, P. Wang, and D. Niyato, "A Neural Network Based Spectrum Prediction Scheme for Cognitive Radio," *Proc. IEEE ICC*, May 2010.

[19] L. Giupponi and A. I. Perez-Neira, "Fuzzy-Based Spectrum Handoff in Cognitive Radio Networks," *Proc. CrownCom*, May 2008.

[20] N. Sharma and A. S. Madhukumar, "Genetic Algorithm Aided Proportional Fair Resource Allocation in Multicast OFDM Systems," *IEEE Trans. Broadcast.*, vol. 61, no. 1, Mar. 2015.

BIOGRAPHIES

CHUNXIAO JIANG [S'09-M'13-SM'15] received the B.S. in information engineering from Beihang University in June 2008, and the Ph.D. in electronic engineering from Tsinghua University in January 2013, both with the highest honors. From February 2013 to June 2016, Dr. Jiang was a postdoc in the Department of Electronic Engineering, Tsinghua University, during which he visited the University of Maryland College Park and the University of Southampton. He is a recipient of the Best Paper Award from IEEE Globecom 2013 and the Best Student Paper Award from IEEE GlobalSIP 2015. Dr. Jiang became a Senior Member of IEEE in 2015. Currently, he is a research-track faculty in the Tsinghua Space Center.

HAIJUN ZHANG (M'13) is a full professor at the University of Science and Technology Beijing, China. From 2014 to 2016, he was a postdoctoral research fellow in the Department of Electrical and Computer Engineering, the University of British Columbia (UBC), Vancouver, Canada. He received his Ph.D. degree from Beijing University of Posts Telecommunications (BUPT). From September 2011 to September 2012, he visit-

ed the Centre for Telecommunications Research, King's College London, London, UK, as a visiting research associate. Dr. Zhang has published more than 70 papers and authored two books. He serves as an editor of the *Journal of Network and Computer Applications*, *Wireless Networks*, *Telecommunication Systems*, and *KSII Transactions on Internet and Information Systems*. He is serving or has served as the leading guest editor for *IEEE Communications Magazine*, *IEEE Transactions on Emerging Topics in Computing*, and *ACM/Springer Mobile Networks & Applications*. He is serving or has served as general co-chair of the 6th International Conference on Game Theory for Networks (GameNets'16) and 5GWN'17, the symposium chair of the GameNets'14, track chair of the 15th IEEE International Conference on Scalable Computing and Communications (ScalCom2015), and co-chair of the Workshop on 5G Ultra Dense Networks in ICC 2017.

YONG REN [SM'16] received his B.S., M.S., and Ph.D. degrees in electronic engineering from Harbin Institute of Technology, China, in 1984, 1987, and 1994, respectively. He worked as a post doctor in the Department of Electronics Engineering, Tsinghua University, China from 1995 to 1997. He is currently a professor in the Department of Electronics Engineering and the director of the Complexity Engineered Systems Lab (CESL) at Tsinghua University. He holds 12 patents, and has authored or co-authored more than 100 technical papers on the behavior of computer network, P2P networks, and cognitive networks. His current research interests include complex systems theory and its applications to the optimization and information sharing of the Internet, Internet of Things and ubiquitous networks, cognitive networks, and cyber-physical systems.

ZHU HAN [S'01-M'04-SM'09-F'14] received the B.S. degree in electronic engineering from Tsinghua University in 1997, and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Maryland, College Park, in 1999 and 2003, respectively. From 2000 to 2002 he was an R&D engineer at JDSU, Germantown, Maryland. From 2003 to 2006 he was a research associate at the University of Maryland. From 2006 to 2008 he was an assistant professor at Boise State University, Idaho. Currently, he is a professor in the Electrical and Computer Engineering Department, as well as in the Computer Science Department at the University of Houston, Texas. His research interests include wireless resource allocation and management, wireless communications and networking, game theory, wireless multimedia, security, and smart grid communication. He received an NSF Career Award in 2010, the Fred W. Ellersick Prize of the IEEE Communication Society in 2011, the EURASIP Best Paper Award for the *Journal on Advances in*

Signal Processing in 2015, and several best paper awards at IEEE conferences. Currently, he is an IEEE Communications Society Distinguished Lecturer.

KWANG-CHENG CHEN [M'89-SM'94-F'07] is a professor with the Department of Electrical Engineering, University of South Florida, after an academic career in Taiwan and an industrial career in the U.S. He has contributed essential technology to various IEEE 802, Bluetooth, and LTE and LTE-A wireless standards. In addition to service with IEEE journals and conferences, he founded and then chairs the TC on Social Networks of the IEEE Communications Society. Dr. Chen is an IEEE Fellow and has received a number of awards, such as the 2011 IEEE ComSoc WTC Recognition Award, the 2014 IEEE Jack Neubauer Memorial Award, and the 2014 IEEE ComSoc AP Outstanding Paper Award. His recent research interests include wireless networks, social networks and network science, cybersecurity, and data analytics.

LAJOS HANZO (<http://www-mobile.ecs.soton.ac.uk>) FREng, FIEEE, FIET, Fellow of EURASIP, DSc, received his degree in electronics in 1976 and his doctorate in 1983. In 2009 he was awarded an honorary doctorate by the Technical University of Budapest, and in 2015 by the University of Edinburgh. During his 38-year career in telecommunications he has held various research and academic posts in Hungary, Germany, and the UK. Since 1986 he has been with the School of Electronics and Computer Science, University of Southampton, UK, where he holds the chair in telecommunications. He has successfully supervised approximately 100 Ph.D. students, co-authored 20 John Wiley/IEEE Press books on mobile radio communications totalling in excess of 10,000 pages, published 1500+ research entries in IEEE Xplore, acted both as TPC and general chair of IEEE conferences, presented keynote lectures, and has been awarded a number of distinctions. Currently he is directing a 60-member academic research team, working on a range of research projects in the field of wireless multimedia communications sponsored by industry, the Engineering and Physical Sciences Research Council (EPSRC) UK, the European Research Council's Advanced Fellow Grant, and the Royal Society's Wolfson Research Merit Award. He is an enthusiastic supporter of industrial and academic liaison and he offers a range of industrial courses. He is also a governor of the IEEE VTS. From 2008 to 2012 he was the editor-in-chief of the IEEE Press and a chaired professor at Tsinghua University, Beijing. His research is funded by the European Research Council's Senior Research Fellow Grant. For further information on research in progress and associated publications, please refer to <http://www-mobile.ecs.soton.ac.uk> Lajos has 24 000+ citations.