Machine Learning Techniques and A Case Study for Intelligent Wireless Networks

Helin Yang, Xianzhong Xie, and Michel Kadoch

Abstract

With the widespread deployment of wireless technologies and IoT, 5G wireless networks will support various communication connectivity and services for the huge number of wireless smart/intelligent devices and machines. The challenge lies in assisting wireless networks to intelligently learn experience, autonomously optimize network configurations and smartly make decisions to support massive wireless smart devices with minimum human intervention, so the diverse and colorful service requirements can be satisfied with the optimum performance. Machine learning, as one of the powerful artificial intelligence tools, is capable of efficiently supporting wireless smart devices by assisting them to smartly observe the environment, analyze data and make decisions with the intelligence. Hence, in this article, we briefly review the major concepts of common machine learning techniques and present their potential applications in intelligent wireless networks, including spectrum sensing, channel estimation, device clustering, behavior prediction, position tracking, data demission reduction, adaptive routing, energy harvesting/efficiency, resource management, and so on. Furthermore, we propose deep reinforcement learning for intelligent resource management in intelligent wireless networks in an exemplary case study. Simulation results demonstrate the effectiveness and advance of machine learning in intelligent wireless networks.

Introduction

With the various application services of devices, things and machines (e.g., mobile phones, vehicles, sensors and industrial machines) in wireless networks, the family of wireless communication and networking technologies has been emerging as a promising vision for fifth generation (5G) wireless networks through realizing industrial and factory automation [1–3]. With these advanced technologies, wireless networks are capable of interconnecting a large number of smart devices (potentially on the order of tens of billions) with the intelligent and reconfigure ability, and smartly make decisions by itself with a "brain" for high-level intelligence [1]. Hence, it has a myriad of applications in various domains, such as smart home/city/grid, e-health, intelligent transportation, automatic industry, meter auto reporting, remote sensing, and so on, which greatly improve the quality of our lives.

Despite the above distinct benefits, facilitating and implementing intelligent wireless networks gives rise to several key challenges. First, considering the large amount of data generated by the huge number of smart devices, the applications of intelligent wireless networks face the challenges of collecting, accessing and processing the massive amount of data, as well as to exploit and analyze the big amount of data toward the behavior and characteristics discovery of wireless networks [1, 4, 5]. Moreover, due to the extreme range of service requirements of wireless smart devices and the complex/dynamic environments, its applications are still not smart enough to tackle optimized physical layer designs, sophisticated learning, complicated decision making and efficient resource management tasks in future wireless networks [1]. To fulfill the potential benefits of intelligent wireless networks and deal with the growing challenges, recent trends in research on machine learning have drawn attention as a promising solution.

Machine learning, as one of the most powerful artificial intelligence tools, has already been widely applied in computer vision, signal/language processing, social behavior analysis, projection management, and so on [6]. Explicitly, it uses statistical techniques to analyze observations/data/experience by finding the patterns and underlying structures, in order to give devices the ability to "learn" automatically without human intervention and adjust actions accordingly. Machine learning mainly consists of three categories: supervised learning, unsupervised learning and reinforcement learning [6]. Supervised learning algorithms are provided with labeled training samples, while unsupervised learning algorithms are not provided with labels (i.e., no output variables). Reinforcement learning algorithms learn how to map situations to actions to maximize a reward by interacting with its environment.

As shown in Fig. 1, thanks to the application of machine learning, the intelligent wireless network is capable of smartly tackling the detection and sensing tasks (e.g., robust detection, efficient data collection), data analysis and discovery tasks (e.g., knowledge discovery, behavior prediction) as well as decision making tasks (e.g., resource management, policy control) from the physical layer to the application layer in intelligent wireless networks. In itself, machine learning offers a versatile set of algorithms to analyze numerous data/observations and discover the depth knowledge. This effectively assists intelligent wireless networks intelligently adapt network protocols and decision

Digital Object Identifier:
10.1109/MNET.001.1900351

Helin Yang is with Nanyang Technological University; Xianzhong Xie is with Chongqing University of Posts and Telecommunications; Michel Kadoch is with the University of Quebec.

0890-8044/19/$25.00 © 2019 IEEE

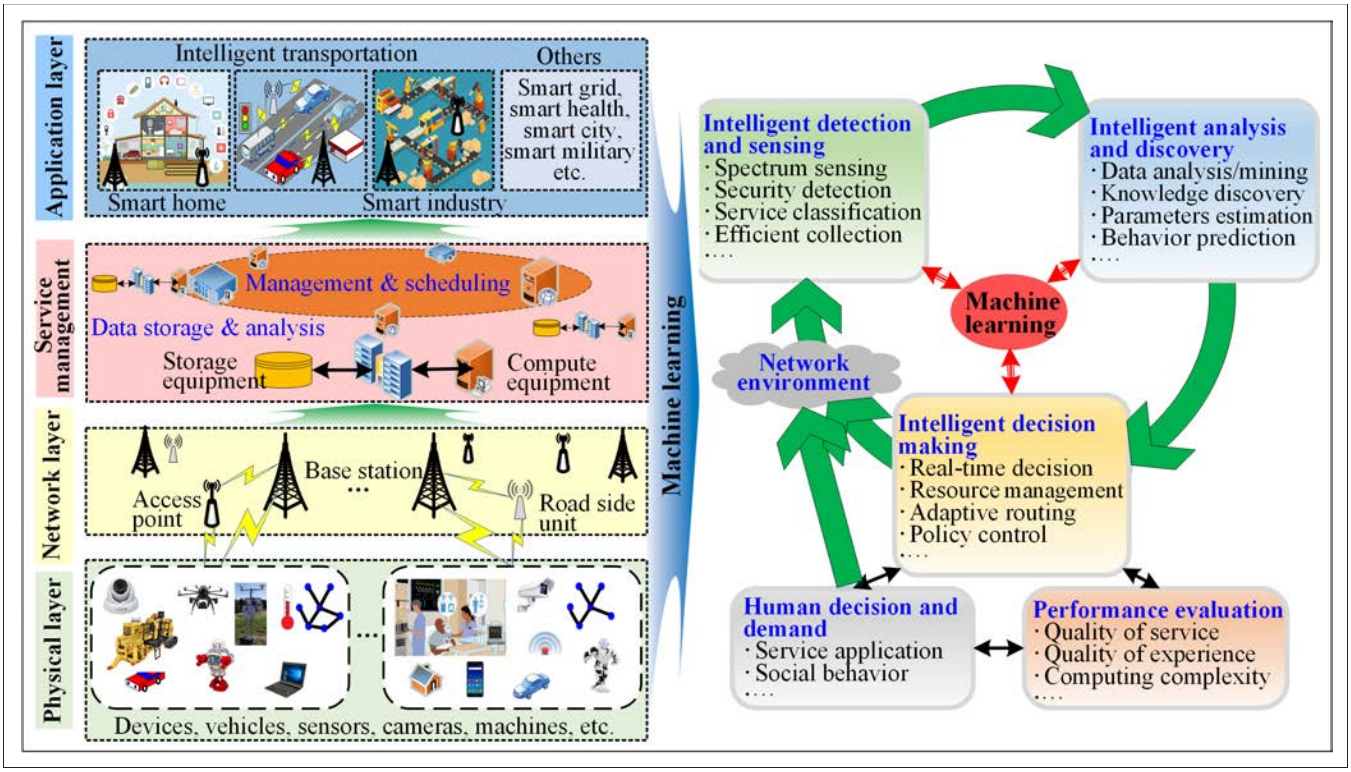

FIGURE 1. Functional diagram of intelligent wireless networks based on machine learning.

Category	Tasks	Algorithms	Applications and References
Supervised learning	Classification	Support vector machine, K-nearest neighbors	Security/interference detection, image/service behavior classification, spectrum sensing [2, 4, 9, 10]
	Regression	Linear regression, support vector regression, Gaussian process regression	Channel estimation, mobility prediction, cross-layer handover [9, 10, 12]
Unsupervised learning	Clustering	K-means clustering, neural network	Device clustering, filtering designs, localization, service segmentation [5, 10, 13, 14]
	Dimension reduction	Principal component analysis, isometric mapping	Big data visualization, interference filtering, data compression, feature elicitation [4, 10, 11, 14]
Reinforcement learning	Policy/value iteration learning	Markov decision process, Q-learning, policy gradient, actor critic, deep Q-network,	Decision making, packet transmission, spectrum access, network association, energy harvesting/efficiency, adaptive routing, resource management [1, 7–11, 14, 15]

TABLE 1. Machine learning techniques and applications in intelligent wireless networks.

making for different services in different complex scenarios, and solves various technical problems, such as signal processing, parameter optimization, behavior analysis, mobile management and resource management [1, 4, 5, 7-15]. However, how to adapt and exploit the family of machine learning algorithms to address the above mentioned problems in intelligent wireless networks remains a significant challenge.

In this article, the goal is to pay more attention to the research on applying machine learning techniques to solve the key challenges in intelligent wireless networks. Table I presents the family-tree of the three categories of machine learning (i.e., supervised, unsupervised and reinforcement learning) and their potential applications in intelligent wireless networks, where each of the following sections will introduce the basic algorithms of one category of machine learning and then discuss several typical examples of applying such algorithms for intelligent wireless networks. After that, we present an exemplary case study on intelligent resource management based on deep reinforcement learning in intelligent wireless networks and evaluate the performance improvement of transmission scheduling accordingly. Finally, we conclude the article.

Supervised Learning for Parameter Estimation

In supervised learning, a set of labeled features or data is used to build the learning model. The model uses the training data to learn the relation

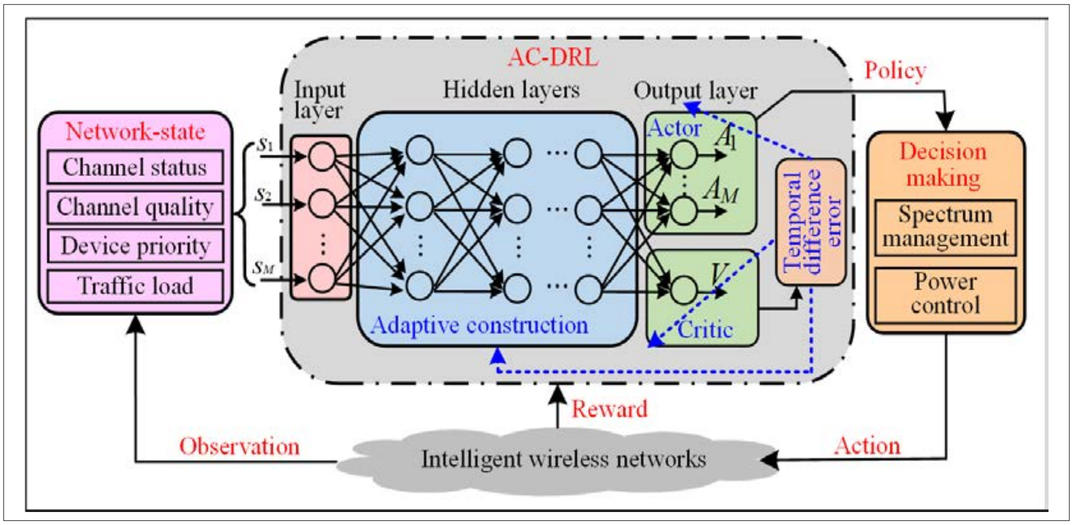

FIGURE 2. The AC-DRL learning framework for intelligent resource management in intelligent wireless networks.

from the inputs to the outputs. Supervised learning tasks can be broadly divided into two subsection: classification and regression.

Classification Models for Spectrum Detection and Behavior Classification

Models: Classification analysis is the task of estimating the relation between the inputs and the discrete outputs, and the output variables are also called categories or labels. A mapping (classifier) function is built by analyzing the input training data in the learning step, and the mapping function is adopted to predict the categorical class labels in the classification step. The support vector machine (SVM) and K-nearest neighbors (KNN) classifiers are most used for classification [6].

SVM is a non-probabilistic and liner/nonlinear classifier that tries to search a linearly separable hyperplane which separates the training data by maximizing the margin with the minimum classification errors. When the training data are non-linear, SVM constructs the data into a higher dimensional feature space by using the kernel trick, and performs as a non-linear classifier.

KNN is a non-parametric tool used for classification. The learning model in KNN is built by finding the most similar points (closest neighbors) in the training samples, and makes an educated guess according to the classification.

Applications: The classification algorithms can be applied for spectrum sensing, security and intrusion detection, interference detection and image classification for intelligent wireless networks [2, 4]. For instance, in intelligent wireless networks, when a large number of smart devices aim to access the spectrum radio, the various channel sensing processes result in high dimensional search problems. In this case, the mentioned SVM and KNN learning models can be applied to detect the channel working status by categorizing each feature vector (channel) into either of the two classes, namely, the "channel idle class" and "channel buy class," which is capable of implicitly/adaptively learning the surrounding environment in an online fashion.

In addition, with the huge diversity of devices' activities (sleep/active or mobility) and services (spectrum access) in intelligent wireless networks, it is hard for all smart devices to follow the same rules and standards in the complex and dynamic environments. The mentioned classification tools constitute powerful data processing techniques which are devised to classify all the possible behaviors, where similar behaviors can be merged in the same group [9], [10]. Such classification tools have the ability to identify the behavior features of devices, and automatically establish the learning models to classify the various activities, which can improve the recognizing ability and smartly establish communication rules among smart devices.

Regression Models for Channel Estimation and Mobility Prediction

Models: Regression analysis aims to estimate or predict continuous quantities. Regression relies on input statistical features to establish the relationship between two or more independent variables. Some typical regression algorithms include linear regression, support vector regression (SVR) and Gaussian process regression [6].

The linear regression model establishes a relationship between one independent variable and one or more other dependent variables by searching the best fit regression line.

SVR has the same main principles as SVM with some minor differences. Instead of minimizing the training error in SVM, SVR tries to minimize the generalization error bound by building the linear/non-linear regression function in a high dimensional feature space.

Gaussian process regression (GPR) is a non-parametric learning tool by undertaking the non-parametric regression with Gaussian processes. As the GPR model is probabilistic, it is powerful to provide uncertainty estimation and to learn smoothness parameters.

Applications: The above mentioned regression models can be used to estimate the channel parameters in intelligent wireless networks. For example, the wireless channel varies rapidly in high mobility scenarios (e.g., vehicular networks), which directly affects the communication performance and quality-of-service (QoS) due to the fast time-varying/multipath fast fading channel response. SVR can be applied to estimate the channel variation by utilizing the prior knower of similar features to identify future unknown changes [12]. In addition, for the non-linear deep fading channel, the non-linear SVR model has the ability to track the channel response by overcoming the unknown and complex estimation difficulties.

Regression models are also widely applied to solve the devices' mobility prediction and cross-layer handover optimization problems for intelligent wireless networks [4, 9, 10]. In large and heterogeneous vehicular intelligent wireless networks, the ability to accurately predict intelligent vehicles' mobility has many enjoyable applications, such as vehicle routing, mobility prediction and congestion avoidance. SVR and Gaussian process regression models can predict the long-range and short-range multimedia traffic load and network overhead with the high accuracy by proposing self-similar covariance functions or learning the non-linear relationships between historical and future features. In addition to traffic prediction, the devices' behavior prediction (e.g., channel access) by using the regression models can effectively solve the handover problem in large-scale intelligent wireless networks.

Unsupervised Learning for Feature or Data Analysis

Different from supervised learning with an amount of labeled data, unsupervised learning is not provided with labels (no output vectors). The objective of unsupervised learning is to analyze the data structure and extract the useful information from the training data without any guidance of an explicit data of interest. It mainly has two subfields: clustering and dimensionality reduction.

Clustering Models for Device Clustering and Localization

Models: Clustering analysis seeks to divide a set of objects into different groups such that the objects of each group are as similar as possible to one another, and different groups are as dissimilar as possible from one another. K-means clustering (KNN) and neural network (NN) are the common clustering algorithms [6].

In the KNN model, the goal is to identify the best K cluster centers and recognize the unlabeled inputs into a given number of groups (classes) in an iterative manner, where each input belongs to the group with the nearest mean.

NN clustering analysis is a powerful tool for solving the pattern recognition problems inspired by the biological brains, where the input data are handled with the non-linear correlations in the hidden layers and the weights in hidden layers finally determine the cluster reconstruction requirements.

Applications: The above mentioned clustering models are capable of solving the clustering problems in intelligent wireless networks, especially in dense devices environments. For example, in order to complete the massive spectrum access services of smart devices in future wireless networks, clustering algorithms can be used to cluster the numerous number of devices into different groups according to their interests [5, 14], which significantly avoids interference, reduces the collision probability and enhances the successful access probability. In addition to the above benefits, a cluster header (CH) is selected to complete the transmission scheduling and data gathering from all the devices in the cluster and send the assembled data to base stations (BSs), which is easy for the transmission scheduling and reduces the energy consumption.

Different from supervised learning with an amount of labeled data, unsupervised learning is not provided with labels (no output vectors). The objective of unsupervised learning is to analyze the data structure and extract the useful information from the training data without any guidance of an explicit data of interest. It mainly has two subfields: clustering and dimensionality reduction.

Moreover, the clustering algorithms are readily invoked for outdoor/indoor localization and tracking in intelligent wireless networks. To address the range-based multi-device positioning and tracking problems, clustering algorithms are able to optimize the clustering processing of massive location data and filter out the extreme positioning references, and then calculate the initial cluster center through analyzing the density of each measurement point. Finally, the position of each device is calculated according to the predictive location and the measurement point sets [10, 13].

Dimensionality Reduction Models for Data Compression and Interference Filtering

Models: Dimensionality reduction seeks to transform a high dimensional data space into a low dimensional structure without losing the useful information of the original data. Some classic dimension reduction algorithms include principal component analysis (PCA) and isometric mapping (ISOMAP) [6].

PCA orthogonally transforms a set of possibly high-dimensional and correlated variables into a lower-dimensional linear subspace with a linear mapping, where the variables are uncorrelated and the variance of the variables is maximized.

ISOMAP is a non-linear dimensionality reduction algorithm. It begins by creating a neighborhood network and then estimates the geodesic distance based on the graph distance between all data points. After that, by decomposing the eigenvalues of the geodesic distance matrix, the low dimensional embedding of the data points can be achieved.

Applications: In intelligent wireless networks with massive training samples, transmitted data packets and application services [4, 10, 11], the above mentioned demission reduction algorithms can efficiently reduce the amount of data by finding a small set of useful variables of the original data, which dramatically decreases the computing time, storage space and model complexity. The transmitted data or packets are aggregated with the transmission cluster heads before being sent to BSs by using PCA and ISOMAP, intern reduces the communication costs and energy consumption in cluster-based intelligent wireless networks.

> Considering that there are uncertainties in the state of the mobile devices and their actions effect the state dynamics, such as the partial observation of the environment and imperfect position tracking/navigation, the decision making problem can be formulated as a POMDP model under the partial knowledge.

In addition, demission reduction algorithms have the ability to separate the desired signal and the noise (or interference) subspace by taking the dimension reduction process [11, 14], which can significantly decrease the additive noise on the channel estimation in intelligent wireless networks. As the input variables of each dimension may be correlated and some dimensions are also mixed with the noise and interference data, which directly degrades the network performance if those useless and interference data are not filtered properly. After filtering the interference data or useless training data, the low dimensional data used in the learning models can greatly improve the positioning accuracy, parameter estimation performance, device behavior prediction, and spectrum sensing accuracy.

Reinforcement Learning for Decision Making

Reinforcement learning (RL) refers to a key type of machine learning where the agent makes decisions on what actions to take in a certain environment, in order to maximize some notions of the cumulative reward. From now on, various algorithms are used to solve the RL problems, such as Markov decision process (MDP), Q-learning, policy gradient, actor critic (AC), and deep reinforcement learning (DRL) [6].

MDP Models for Network Association and Vehicular Routing

Models: MDP provides a mathematical representation of the decision making process, and in the RL formwork, the optimal policy searching process can be modeled as a MDP. The agents interact with the environment in discrete time steps. At each time, the agent takes the action a_t from the current state s_t to a new state s_{t+1}, and calculates the corresponding reward U_t. During this process, the probability of moving the current state into a new state is described by the transition probability $P(s_{t+1} | s_t, a_t)$. The environment evaluates the quality of the policy based on the immediate reward as well as its cumulative reward, and explores the optimal policy in the next time step. In MDP, the future states are determined by the current state and action rather than the former ones, and future states only depend on the current state, and the transition probability can be conditionally independent, which guarantees the Markov property. MDP assumes that the state is known when each action is to be taken where the environment is fully observable; otherwise the policy cannot be achieved. When the network knows the partial knowledge, the problem is viewed as a partially observable Markov decision process (POMDP).

Applications: In the context of intelligent wireless networks, the MDP/POMDP models can be applied to solve the decision making problems [4, 10, 11, 14] (e.g., spectrum access, network association, energy harvesting and load-balancing), where the smart devices can be regarded as agents and the network constitutes the environment. For example, in intelligent wireless networks, a large amount of wireless smart devices always choose the evolved NodeB (eNB) with the best-signal-quality for the attachment, thereby leading to serious network congestion and overload. For this case, the eNB selection problem can be formulated as a MDP/POMDP, where the fixed bandwidth of eNB, the limited energy of devices and the time-variant channels are defined as the environment, and the devices' connecting selection, their transmission power levels and the number of transmission packets are regarded as the actions.

Considering that there are uncertainties in the state of the mobile devices and their actions effect the state dynamics, such as the partial observation of the environment and imperfect position tracking/navigation, the decision making problem can be formulated as a POMDP model under the partial knowledge. For instance, in large-scale vehicular wireless networks, the traffic situations are highly complex, uncertain, dynamic and only partially observable; hence, the decision making problem (e.g., automated driving, adaptive routing) can be formulated as a POMDP model. Adopting POMDP in vehicular networks can effectively avoid collisions and decrease the traffic congestion, enhance the driving safety and vehicle-network resource utilization efficiency [11].

Value/Policy-Iteration Learning Models for Policy Control

Models: The RL algorithms can be divided into two groups: value iteration and policy iteration. Value iteration (e.g., Q-learning) starts with a random value function and then iteratively updates the value function until achieving the optimal value function. The best policy can be derived based on the optimal value function. By contrast, in policy-iteration (e.g., policy gradient), you randomly select a starting policy and iterate toward the optimal solution until the policy converges by finding the value function of that policy [6].

Q-learning is a model-free and value-iteration RL algorithm, which solves the MDP problem in an unknown environment. The agent in the Q-learning model uses a Q function to estimate its accumulated reward.

Policy Gradients (PG) seeks to directly optimize a policy function (instead of a Q function in Q-learning) in the policy space. The optimized approximation policy is learned by directly maximizing the expected reward by adopting the gradient methods.

AC combines the benefits of the value-iteration and policy-iteration models, which consists of the actor and the critic, where the actor is represented through adopting a control policy with action selections and the critic evaluates the input policy by a reward function.

DRL applies the deep learning techniques (e.g., deep neural networks) within RL by directly using the deep learning network to represent the value function or policy model, such as the deep Q-learning and deep Q-network.

Applications: The above mentioned RL algorithms have been widely applied in large scale

Parameters	Values	Parameters	Values
Number of devices	1000, 500, …, 3500	Background noise power	−114dBm
Channel model	Frequency selective fading	Number of time slots	5000
Packet size	2000 bits	Number of RBs	16
Buffer size	10 packets	Number of hidden layers	3
Time slot duration	10 ms	Learning rate of NN	0.02
Packet arrival rate	0.01/(10 ms)	Training error accuracy	1×10^{-4}
Device power consumption in “active” status	35 mW	Discount factor	0.002
Device power consumption in “sleep” status	1 mW		

TABLE 2. Simulation parameters.

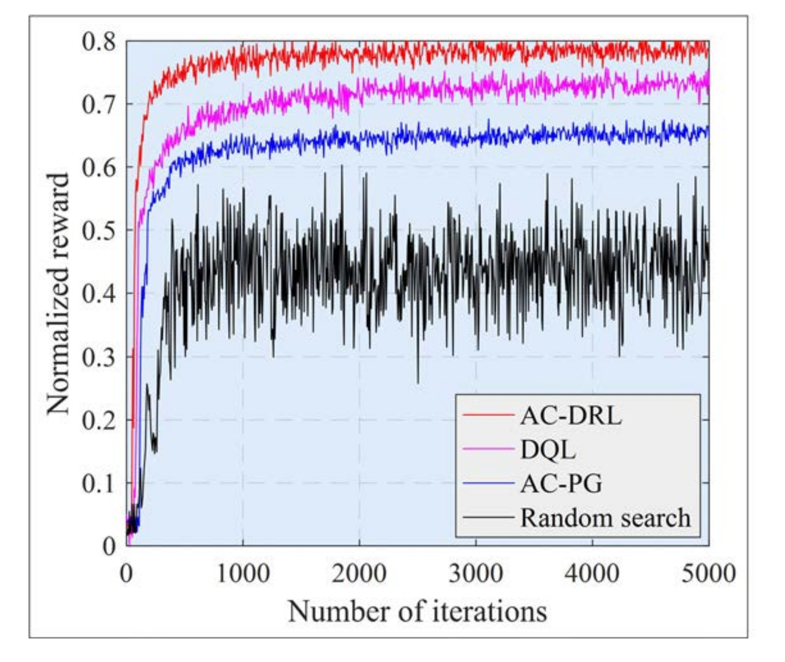

FIGURE 3. Learning process for the four approaches.

intelligent wireless networks, for supporting the intelligent decision making [1, 7–11, 14, 15], such as resource management, channel access, interference coordination, transmission scheduling, power control and so on. With the help of the RL algorithms, the network has the ability to smartly manage its inter-operation and make decisions independently among devices with minimal human interaction, which makes the network with high-level intelligence. For example, a large number of smart devices entails a significant increase in energy consumption in intelligent wireless networks, so the energy minimization problem becomes more challenging. Hence, the scheduling framework incorporating RL enables the scheduler to intelligently develop an association between the optimal action and the current state of the environment to minimize the energy consumption with variability of workloads.

The conventional RL (Q-learning or AC) is suitable to make decisions with handcrafted features or low-dimensional data, while DRL is able to learn their action-value policies directly from complex high-dimensional inputs. For instance, in dynamic networks, the channel conditions, devices' requirements, and caches' storage are all dynamically changing, the network has a large number of environment sates (e.g., devices status (sleep or active), channel quality and channel status (busy or idle)), DRL has the ability to solve the complex and large state-space resource allocation problem by using deep learning algorithms instead of estimating the value function and it is proven to be more advantageous and more robust learning [14].

Case Study: Deep Reinforcement Learning for Resource Management

This section presents an actor-critic deep reinforcement learning approach (called AC-DRL) for intelligent resource management in intelligent wireless networks, as shown in Fig. 2. In intelligent wireless networks, all smart devices can be regarded as agents and the network constitutes the environment. First, each device intelligently observes its current network state (e.g., channel status (busy or idle), channel quality, devices priority and traffic load) by integrating with the environment. Then, it makes a decision and selects an action by itself based on its learned policy strategy in a decentralized way. After that, the environment provides a new network state and an immediate reward. According to the feedback, all devices smartly learn a new policy in the next step. The optimum parameters of both the actor part and the critic part can be achieved with an infinite number of learning steps, when the AC-DRL learning converges to the optimum value function and policy. Finally, the best actions for intelligent resource management are achieved in intelligent wireless networks. Compared with centralized intelligent wireless networks, the advantage of decentralized AC-DRL learning is that each device can learn independently based on its local observation information, rather than continuously exchange information among devices.

The input of the AC-DRL framework is the network-state vector $\boldsymbol{s} = [s_1, ... s_M]^T$ with the number of M network states. The output of the AC-DRL framework is the estimated functions of both the actor function vectors $\boldsymbol{A} = [A_1, ..., A_M]^T$ and the critic function V. The multiple hidden layers perform computations on the weighted inputs (network states) and produce net input, which is then applied with activation functions to produce the actual output (functions of both the actor and the critic). In intelligent wireless networks, the optimal policy and intelligent decision making (spectrum access, spectrum handoff and transmission power control) can be carried out to support the network services' requirements in the AC-DRL framework, and it provides considerable actions to the physical environment. Generally, the learning mechanism is driven by the reward. In our proposed leaning framework, the expected reward considers the successful packet transmission probability, the power consumption and the blocking probability in the intelligent wireless network [14, 15]. The policy used in the intelli-

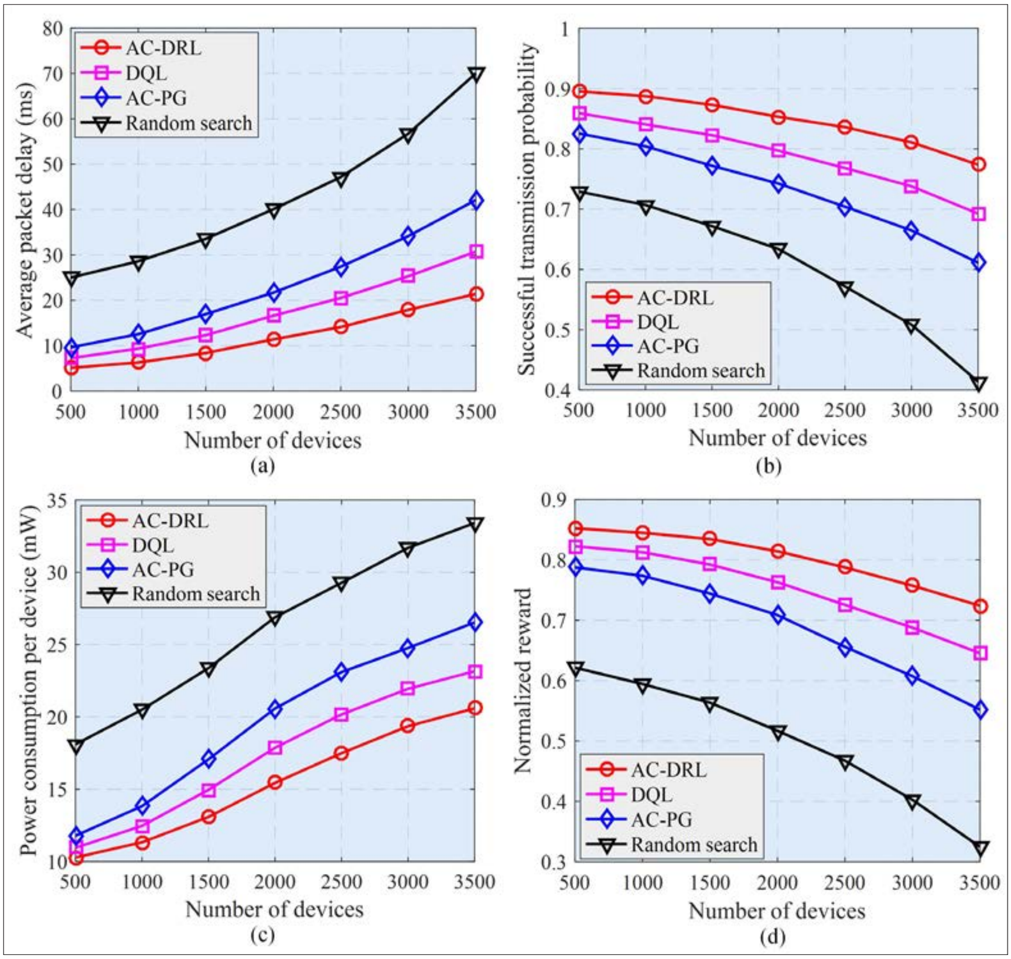

FIGURE 4. Performance comparison in terms of a) average delay per packet; b) successful transmission probability; c) average power consumption per device; d) reward versus a number of smart devices.

gent wireless network for spectrum management and power control is random at the beginning and then gradually improved with the updated AC-DRL framework.

Performance Analysis: We evaluate the performance of our proposed AC-DRL approach, and compare it with the following approaches: 1. Classical AC approach based on PG (denoted as AC-PG); 2. Deep Q-network approach (denoted as DQN); 3. Random search approach, (denoted as random search). We consider that the devices are randomly distributed in a circular cell area with a radius of 500m; the packet generated by each device forms a Poisson process. The devices are divided into two priority levels, where the resource blocks (RBs) first are allocated to the devices with the higher priority level [15]. The main simulation parameters are listed in Table 2 [11, 15].

Figure 3 shows the learning process of the four approaches in terms of the reward performance when the number of smart devices is 2500. We can see that the three RL approaches greatly outperform the random search approach; specifically, the proposed AC-DRL approach achieves the best reward performance with the fastest convergence rate compared with the other three approaches. For the DQN approach, it needs to search the Q function approximator, which may fail miserably with the huge number of devices. In addition, the AC-PG learning approach has a fast convergent rate, but it may converge to the local optimal point. For the random search approach, its performance is the worst among the four approaches, because it randomly searches the policy only based on the current immediate reward, but it has the simple control structure. Our proposed approach adopts the deep learning network to approximate both the actor function and critic function, and the optimal policy will be learned after a finite number of learning steps.

Figure 4 shows the performance comparison for the four approaches with the range of the number of devices in intelligent wireless networks. The network resource is limited and fixed, when a large number of devices' packets need to be transmitted as the increase in devices, which results in the frequent handover process, blocking and retransmission, all these factors increase the average packet delay and decrease the successful transmission probability, thus the lower reward is obtained in the large number of devices regions. In addition, the high frequent handover and retransmission increase the extra power consumption. However, the RL approaches significantly outperform the random search approach by searching the best transmission scheduling policy, especially in the large number of devices regions. Moreover, the proposed approach achieves the best performance among them, which proves that it is a great intelligent learning approach to deploy the intelligent resource management for intelligent wireless networks.

The RL approaches significantly outperform the random search approach by searching the best transmission scheduling policy, especially in the large number of devices regions. Moreover, the proposed approach achieves the best performance among them.

Conclusions

The key challenge for future wireless networks is how to intelligently support the huge number of wireless smart devices and machines under diverse service requirements. In this article, we have briey provided a comprehensive survey of the major families of machine learning algorithms and discussed their potential applications in the context of intelligent wireless networks, in order to facilitate future networks with high-level intelligence. Furthermore, an exemplary case study and simulation analysis of intelligent resource management are provided to demonstrate the advantage and significance of machine learning in intelligent wireless networks. In a nutshell, machine-learning-based physical layer design, decision making, network management and resource optimization is an exciting area for future intelligent wireless networks.

Acknowledgment

This work was supported by the National Nature Science Foundation of China under Grant no. 61502067; the Key Research Project of the Chongqing Education Commission under Grant no. KJZD-K201800603; the Key Project of Science and Technology Research of the Chongqing Education Commission under Grant no. KJZD-M201900602; and the Chongqing Nature Science Foundation under Grant no.CSTC2018jcyjAX0432 and CSTC2016jcyjA0455.

References

[1] I. Kakalou *et al.*, "Cognitive Radio Network and Network Service Chaining toward 5G: Challenges and Requirements," *IEEE Commun. Mag.*, vol. 55, no. 11, Nov. 2017, pp. 145–51.

[2] M. Jalil Piran *et al.*, "QoE-Driven Channel Allocation and Handoff Management for Seamless Multimedia in Cognitive 5G Cellular Networks," *IEEE Trans. Vehic. Tech.*, vol. 66, no. 7, Jul. 2017, pp. 6569–85.

[3] X. Zhang, W. Cheng, and H. Zhang, "Heterogeneous Statistical QoS Provisioning over 5G Mobile Wireless Networks," *IEEE Network*, vol. 28, no. 6, Nov. 2014, pp. 46–53.

[4] Q. Mao, F. Hu, and Q. Hao, "Deep Learning for Intelligent Wireless Networks: A Comprehensive Survey," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 4, Fourth Quarter 2018, pp. 2595–2621.

[5] M. Mohammadi and A. Al-Fuqaha, "Enabling Cognitive Smart Cities Using Big Data and Machine Learning: Approaches and Challenges," *IEEE Commun. Mag.*, vol. 56, no. 2, Feb. 2018, pp. 94–101.

[6] E. Alpaydin, *Introduction to Machine Learning*, MIT press, 2014.

[7] K. Zaheer *et al.*, "A Survey of Decision-Theoretic Models for Cognitive Internet of Things (CIoT)," *IEEE Access*, vol. 6, Apr. 2018, pp. 22489–22512.

[8] S. Ayoubi *et al.*, "Machine Learning for Cognitive Network Management," *IEEE Commun. Mag.*, vol. 56, no. 1, Jan. 2018, pp. 158–65.

[9] C. Jiang *et al.*, "Machine Learning Paradigms for Next-Generation Wireless Networks," *IEEE Wireless Commun.*, vol. 24, no. 2, Apr. 2017, pp. 98–105.

[10] M. A. Alsheikh *et al.*, "Machine Learning in Wireless Sensor Networks: Algorithms, Strategies, and Applications," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 4, Apr. 2014, pp. 1996–2018.

[11] H. Ye *et al.*, "Machine Learning for Vehicular Networks: Recent Advances and Application Examples," *IEEE Vehic. Tech. Mag.*, vol. 13, no. 2, June 2018, pp. 94–101.

[12] A. Charrada and A. Samet, "Joint Interpolation for LTE Downlink Channel Estimation in Very High-Mobility Environments with Support Vector Machine Regression," *IET Commun.*, vol. 10, no. 17, Nov. 2016, pp. 2435–44.

[13] T. Wang *et al.*, "Measurement Data Classification Optimization Based on a Novel Evolutionary Kernel Clustering Algorithm for Multi-Target Tracking," *IEEE Sensors J.*, vol. 18, no. 9, May 2018, pp. 3722–33.

[14] Y. He *et al.*, "Software-Defined Networks with Mobile Edge Computing and Caching for Smart Cities: A Big Data Deep Reinforcement Learning Approach," *IEEE Commun. Mag.*, vol. 55, no. 12, Dec. 2017, pp. 31–37.

[15] K. A. M, F. Hu, and S. Kumar, "Intelligent Spectrum Management Based on Transfer Actor-Critic Learning for Rateless Transmissions in Cognitive Radio Networks," *IEEE Trans. Mobile Comput.*, vol. 17, no. 5, May 2018, pp. 1204–15.

Biograhies

Helin Yang [S'15] (hyang013@e.ntu.edu.sg) is a Ph.D. candidate in the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore. He serves as a reviewer for IEEE international journals such as *IEEE Communications Magazine*, *IEEE Wireless Communications Magazine*, *IEEE Transactions on Wireless Communications*, among others. His current research interests include wireless communication, visible light communication, Internet of Things and resource management.

Xianzhong Xie [M'18] (xiexzh@cqupt.edu.cn) received his Ph.D. degree in communication and information systems from Xidian University, China, in 2000. He is currently a professor with the School of Optoelectronic Engineering, and Director of the Chongqing Key Lab of Computer Network and Communication Technology, at Chongqing University of Posts and Telecommunications (CQUPT), China. His research interests include MIMO precoding, cognitive radio networks, and cooperative communications.

Michel Kadoch [S'86, M'91, SM'04] (michel.kadoch@etsmtl.ca) received the Ph.D. degree from Concordia University in 1992. He is currently a full professor with Ecole de Technologie Suprieure (ETS), University of Quebec, Montreal. He is the director of the research laboratory LAGRIT at ETS. He is also an adjunct professor at Concordia University, Canada. As the principal investigator, he has managed and participated actively in a research program on QoS for multicast in high speed networks sponsored by Bell Canada and NSERC. He is presently working on reliable multicast in wireless ad hoc networks and 5G heterogeneous networks.