MACHINE LEARNING FOR VEHICULAR NETWORKS

Recent Advances and Application Examples

Hao Ye, Le Liang, Geoffrey Ye Li, JoonBeom Kim, Lu Lu, and May Wu

The emerging vehicular networks are expected to make everyday vehicular operation safer, greener, and more efficient and pave the path to autonomous driving in the advent of the fifth-generation (5G) cellular system. Machine learning, as a major branch of artificial intelligence, has been recently applied to wireless networks to provide a data-driven approach to solve traditionally challenging problems. In this article, we review recent advances in applying machine learning in vehicular networks and attempt to bring more attention to this upcoming area.

Background and Motivation

With vehicles becoming more aware of their environments and evolving toward full autonomy, a new level of connectivity among them is necessary, leading to the concept of connected vehicles. The emerging vehicular network has been regarded as an important component of the development of the intelligent transportation system (ITS) and smart cities. It is expected to enable a whole new set of applications, ranging from road safety improvement to traffic efficiency optimization, from autonomous driving to ubiquitous Internet access in vehicles [1], [2]. This new generation of networks will ultimately have a profound impact on society and the daily lives of millions of people around the world.

Digital Object Identifier 10.1109/MVT.2018.2811185
Date of publication: 24 April 2018

Despite its significant potential to transform the everyday vehicular experience, the vehicular network also brings unprecedented challenges unseen in traditional wireless communications systems due to the strict and diverse quality of service (QoS) requirements as well as the inherent dynamics in vehicular environments, such as fast-varying wireless propagation channels and ever-changing network topology. To tackle such challenges, various communication standards, e.g., dedicated short-range communications in the United States and the ITS-G5 in Europe, both based on the IEEE 802.11p standard, have been developed across the globe [1]. Recently, the Third Generation Partnership Project (3GPP) has also started projects toward supporting vehicle-to-everything (V2X) services in long-term evolution (LTE) networks and the future 5G cellular system [1], [2].

Meanwhile, with high-performance computing and storage facilities and various advanced onboard sensors, such as lidar, radar, and cameras, vehicles will be more than just a simple means of transportation. They are generating, collecting, storing, processing, and transmitting massive amounts of data used to make driving safer and more convenient, as illustrated in Figure 1. These rich sources of data will necessarily provide new dimensions and abundant opportunities to explore the design of reliable and efficient vehicular networks. However, traditional communication strategies are not meant to handle and exploit such rich information.

Machine learning, as a major branch of artificial intelligence, builds intelligent systems to operate in complicated environments and has found many successful applications in computer vision, natural language processing, and robotics [3], [4]. It develops efficient methods to analyze a huge amount of data by finding patterns and underlying structures, which can be beneficial to supporting future smart radio terminals [5]. Moreover, machine learning represents an effective data-driven approach, making it more robust to handle heterogeneous data because no explicit assumptions are made on the data distribution. Machine learning provides a versatile set of tools to exploit and mine multiple sources of data generated in vehicular networks. This will help the system make more informed and data-driven decisions and alleviate communications challenges as well as provide unconventional services, such as location-based services, real-time traffic flow prediction and control, and autonomous driving. However, how to adapt and exploit such tools to serve the purpose of vehicular networks remains a challenge and represents a promising research direction. The objective of this article is to bring more attention to this emerging field because the research on applying machine learning in vehicular networks is still in its infancy.

Machine-Learning Tools

Machine-learning methods can be roughly divided into three major categories: supervised, unsupervised, and reinforcement learning. Table 1 illustrates the family tree of machine learning with applications in wireless networks. Other learning schemes, such as semisupervised, online, and transfer learning, can be viewed as variants of these three basic types. In general, machine learning

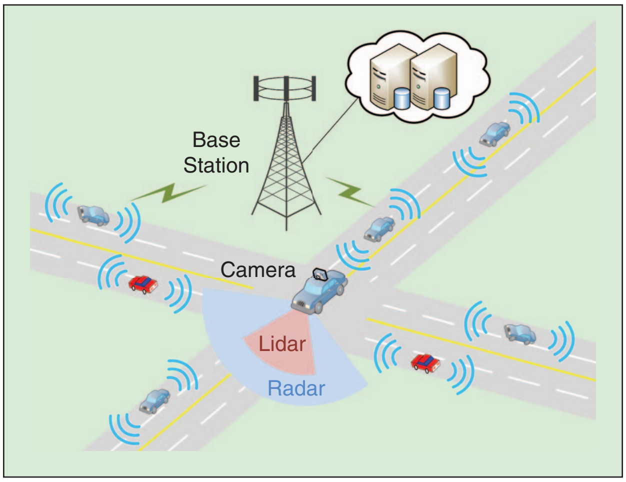

FIGURE 1 An illustrative structure of vehicular networks.

involves two stages, i.e., training and testing. In the training stage, a model is learned based on training data, whereas in the testing stage, the trained model is applied to produce the prediction.

Supervised Learning

Supervised learning receives a labeled data set and can be further divided into classification and regression types. Each training sample comes with a discrete (classification) or continuous (regression) value called a *label* or *ground truth*. The ultimate goal of supervised learning is to gain the mapping from the input feature space to the label or decision space.

Classification algorithms assign a categorical label to each incoming sample. In wireless networks, the classification problems include detecting whether the networks have been intruded or parts of the system are malfunctioning. Some classic algorithms in this category include Bayesian classifiers, k-nearest neighbors, decision trees, support vector machines, and neural networks [3].

Instead of discrete outputs, regression algorithms predict a continuous value corresponding to each sample, such as estimating a house price given its associated feature inputs. In wireless networks, the regression algorithms can potentially be used to predict channel

TABLE 1 ***An overview of machine-learning algorithms and applications in wireless networks.***

Category	*Tasks*	*Algorithms*	*Applications Examples*
Supervised learning	Classification	Neural networks, decision tree, SVM	Intrusion/fault/anomaly detection
	Regression	Logistic regression, SVR, Gaussian process for regression	Throughput prediction, channel parameter regression
Unsupervised learning	Clustering	k means, spectrum clustering	Congestion control, hierarchical routing
	Dimension reduction	Manifold learning, LLE, ISOMAP	Data aggregation
Reinforcement learning	Policy learning	Q learning	Resource management, routing

SVM: support vector machines; SVR: support vector regression; LLE: local linear embedding; ISOMAP: isometric mapping.

parameters, network throughput, and so forth. Classic algorithms include logistic regression, support vector regression, and the Gaussian process for regression [3].

Unsupervised Learning

For supervised learning, with enough data, the error rate can be reduced close to the minimum error rate bound. However, a large amount of labeled data are often hard to obtain in practice. Therefore, learning with unlabeled data, known as *unsupervised learning*, has attracted more attention. This method of learning aims to find efficient representation of the data samples, which might be explained by hidden structures or hidden variables, which can be represented and learned by Bayesian learning methods.

Clustering is a representative problem of unsupervised learning, grouping samples into different clusters depending on their similarities. Input features could be either the absolute description of each sample or the relative similarities between samples. In wireless sensor networks, routing algorithms of hierarchical protocols need to cluster nearby nodes into a group because it is more energy efficient for members within a cluster to send messages to a central node before intercluster transmission. Classic clustering algorithms include k means, hierarchical clustering, spectrum clustering, and the Dirichlet process [3].

Another important class of unsupervised learning is dimension reduction, which projects samples from a high-dimensional space onto a lower one without losing much information. In many scenarios, the raw data come with high dimension, and we may want to reduce the input dimension for various reasons. One is the so-called curse of dimensionality, which describes the problematic phenomenon that arises when the dimension becomes huge. In optimization, clustering, and classification, the model complexity and the number of required training samples dramatically grow with the feature dimension. Another reason is that the inputs of each dimension are usually correlated, and some dimensions may be corrupted with noise and interference, which will degrade the learning performance significantly if not handled properly.

A typical application example in wireless networks is the data aggregation performed by vehicular cluster heads before transmission to infrastructure nodes to reduce communication costs in cluster-based vehicular networks. Some classic dimension reduction algorithms include linear projection methods, such as principal component analysis, and nonlinear projection methods, such as manifold learning, local linear embedding, and isometric mapping [3].

Reinforcement Learning

Reinforcement learning deciphers what to do, i.e., how to map situations to actions, through interacting with the environment in a trial-and-error search so as to maximize a reward, and it comes without explicit supervision. A Markov decision process (MDP) is generally assumed in reinforcement learning, which introduces actions and (delayed) rewards to the Markov process. The learning Q function is a classic model-free learning approach to solve the MDP problem, without the need for any information about the environment. This Q function estimates the expectation of sum reward when taking an action in a given state, and the optimal Q function is the maximum expected sum reward achievable by following any policy of choosing actions. Reinforcement learning can be applied in vehicular networks to handle the temporal variation of wireless environments, which will be discussed in detail in the "Intelligent Wireless Resource Management" section.

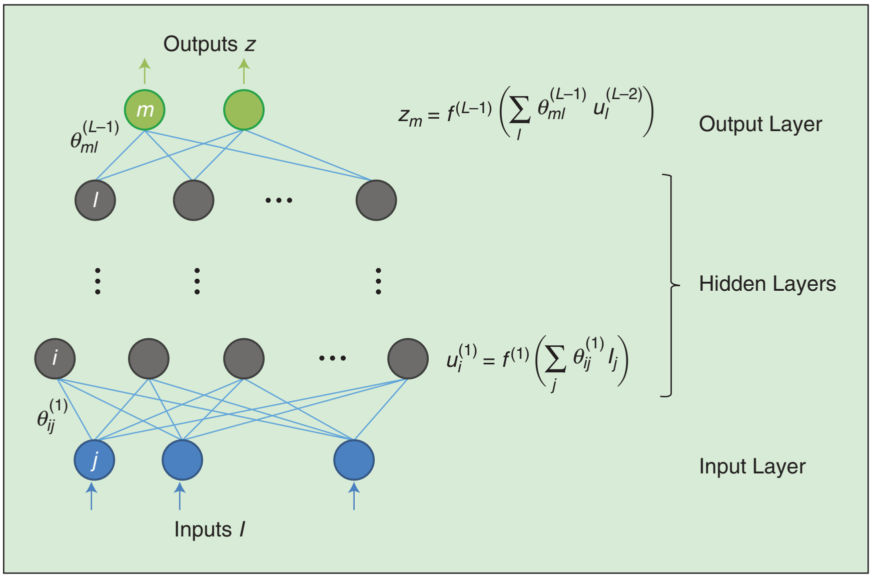

FIGURE 2 An example of deep neural networks.

Deep Learning

Deep learning is a deeper version of neural networks that consists of multiple layers of neurons, as shown in Figure 2. Recently, it has made significant advances on various machine-learning tasks. Deep learning aims to understand the data representations, which can be built in supervised, unsupervised, and reinforcement learning. In Figure 2, the input layer is at the bottom, where each node in the figure represents a dimension of input data. The output layer is at the top, corresponding to the desired outputs, whereas the layers in the middle are called *hidden layers*. Each neuron in the network represents a nonlinear transform, such as a sigmoid function or a rectified linear unit function, on a weighted sum of a subset of neurons in its lower layers. Typically, the number of hidden layers and the number of nodes in each layer are large because the neural network's representation ability grows as the hidden layers become deeper. However, deeper networks bring new challenges, such as needing much more training data and gradients of networks easily exploding or vanishing [4].

With the help of faster computation resources, new training methods (new activation functions, pretraining), and new structures (batch norm, residual networks), training such deep architecture becomes possible. Deep learning has been widely used in such areas as computer vision, speech recognition, and natural language processing and greatly improved state-of-the-art performance in these areas. Depending on applications, different structures can be added to the deep networks, e.g., convolutional networks share weights among spatial dimensions, whereas recurrent neural networks (RNNs) and long short-term memory (LSTM) share weights among the temporal dimensions [4].

Data-Driven Decision Making in Vehicular Networks

To overcome the unprecedented challenges encountered in vehicular networks, it is necessary to rethink traditional approaches to wireless network design, especially given the rich sources of data from various onboard sensors, roadside monitoring facilities, historical transmissions, and so forth. Indeed, it is highly desirable to devise efficient methods to interpret and mine the massive amounts of data and facilitate more data-driven decision making to improve vehicular network performance. Machine learning represents an effective tool to serve such purposes with proven good performance in a wide variety of applications, as demonstrated by some preliminary examples in this section.

Traffic Flow Prediction

Timely and accurate acquisition of traffic flow information is a critical element of ITS deployment because it lays out the foundation for many other services or applications, such as traffic congestion alleviation, fuel consumption reduction, and various location-based services. The objective of traffic flow prediction is to infer from multiple sources of data, including historical and real-time traffic data, the traffic flow information for a wide variety of ITS-related applications. Machine-learning tools can be exploited to produce prediction outputs with accuracy that is hardly achievable using conventional approaches. In [6], a probabilistic graphical model, Poisson regression trees (PRT), has been used for two correlated tasks, LTE connectivity prediction and vehicular traffic prediction. The PRT is used for modeling the count data and is similar to decision trees, where each inner node represents the splitting criterion. Information about the congestion, performance of the communications system, and vehicular traffic information is used to enhance the prediction performance. A novel deep-learning-based traffic flow prediction method based on a stacked autoencoder model has further been proposed in [7], where autoencoders are used as building blocks to represent traffic flow features for prediction and achieve significant performance improvement.

Local Data Storage in Vehicular Networks

The position and connectivity of vehicles are constantly changing in a vehicular environment. However, some data, such as road status and camera sensor information, are region specific and can be used for local traffic information acquisition and estimation, which will be beneficial for load balancing and user-behavior-based adjustment. In vehicular networks, data are naturally generated and stored across different units in the network, e.g., vehicles, roadside units, and remote clouds. A framework has been developed in [8] to store such data in vehicles without any support from the infrastructure. By using unicast transmission to transmit data between vehicles, the region-specific data will always be kept in the region of interest. Initial selection of the next data carrier vehicle node is based on fuzzy logic instant evaluation, which is further refined through applying reinforcement learning. The data carrier node selection takes into account throughput, velocity, and bandwidth efficiency through the fuzzy-logic-based short-term evaluation and also guarantees the long-term rewards through applying Q learning. Reinforcement learning has been further applied to find efficient routing strategy to transfer data from the source node to the selected data carrier node.

Network Congestion Control

In the urban environment, intersections are critical places where congestion of vehicles and communication networks often takes place. A central controlled approach to manage congestion at intersections has been presented by [9] with the help of a specific unsupervised learning algorithm, k-means clustering. The approach basically addresses the congestion problem when vehicles stop at a red light in an intersection, where the roadside infrastructures observe the wireless channels to measure and control channel congestion. Transmission data are clustered into different groups through the use of a k-means clustering mechanism according to their features, including the message size, validity of messages, distances between vehicles and roadside infrastructure, types of messages, and direction of message senders. Each cluster is provided with independent communication parameters, including transmission rate, transmission power, contention window size, i.e., the maximum back-off time, and arbitration interframe spac-

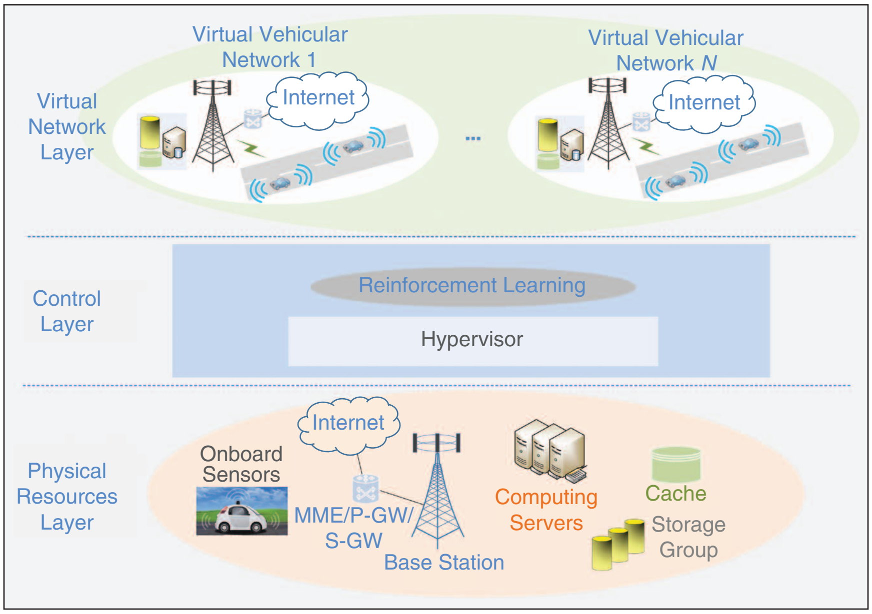

FIGURE 3 An illustration of virtual vehicular networks. MME: mobility management entity; P-GW: packet data network gateway; S-GW: serving gateway.

ing, i.e., the minimum time. The channel has to be unoccupied before transmission so that collision is avoided.

Intelligent Wireless Resource Management

Judicious management of various resources, such as spectrum, transmission power, storage, and computing, is critical for the proper functioning of vehicular networks. Currently, a mainstream approach to resource management is to formulate an optimization problem and then obtain an optimal or suboptimal solution depending on the performance-complexity tradeoffs. However, in practice, vehicular networks are highly dynamic, where the channels and network topology are constantly changing, leading to time-varying optimal solutions. As a result, the optimization problem needs to be recomputed every time a small change occurs in the system, thus incurring huge network overhead. Fortunately, reinforcement learning can serve as an effective alternative solution to the challenge, which learns to interact with the unknown environment, adapts to environmental changes, and takes proper actions.

Load Balancing and Vertical Control

There are potential patterns and regularities of spatial–temporal distribution in traffic flow every day. Reinforcement learning provides an effective tool to utilize such information to address the user association problem in a dynamic vehicular environment. An online reinforcement learning approach for user association with load balancing has been proposed in [10], where an initial association decision is first made using reinforcement learning simply based on the current information. Afterward, the base station keeps accumulating such information and uses historical association patterns to update the association results directly and adaptively. Along this line, a fuzzy-*Q*-learning-based vertical handoff strategy for heterogeneous vehicular networks has been proposed in [11], which determines network connectivity based on four input parameters: received signal strength value, vehicle speed, data quantity, and the number of users associated with the targeted network. The proposed learning-based strategy ensures seamless mobility management without the need for prior knowledge on handoff behavior.

Virtual Resource Allocation

Vehicular clouds, as shown in Figure 3, consisting of various onboard units, roadside units, and remote cloud servers, provide a pool of processing, sensing, storage, and communication resources that are dynamically provisioned for ITS services. How to dynamically allocate the resources to maximize the QoS for end users with small provisioning overhead is a nontrivial task. In [12], the resource allocation problem is modeled as an MDP, where the set of states contains all possible configurations of the allocated resources and the actions are defined as the transition from one state to another. A reinforcement learning framework has been proposed for resource provisioning to cater for such dynamic demands of resources while respecting stringent QoS requirements. A two-stage delay-optimal dynamic virtualization radio scheduling scheme has been developed in [13], which takes joint consideration of the large timescale factors, such as the traffic density, and short-term factors, such as the channel and queue state information. The dynamic delay-optimal virtualization radio resource management has been formulated as a partially observed MDP, which is then solved through an online distributed learning approach.

A unified framework has been proposed in [14] for dynamic orchestration of networking, caching, and computing. The resource allocation problem in the unified framework is formulated as a joint optimization problem. To deal with the high complexity of the joint optimization problem, a deep reinforcement learning approach has been proposed, and desirable performance has been demonstrated. In the future, network slicing will be built on the virtual networks so that the logical network functions and the parameter configurations can be tailored to meet the requirement of a specific service. Learning the slicing management according to the arrival traffic will be essential to support different use cases in vehicular networks.

Distributed Resource Management

There have been many interesting works on resource allocation for device-to-device-based vehicular communications. Most of them are centralized, where the central controller collects information and makes decisions for all of the vehicles. Nevertheless, centralized control schemes will incur huge overheads to acquire the global network information, which grows dramatically with the number of vehicular links. As shown in Figure 4, we have developed a decentralized resource allocation mechanism for vehicular networks based on deep reinforcement learning [14], which is used to find the mapping between the partial observations of each vehicle agent and the optimal resource allocation solution. In particular, this method can address the stringent latency requirement on vehicle-to-vehicle (V2V) links, which is usually hard to deal with using existing optimization approaches.

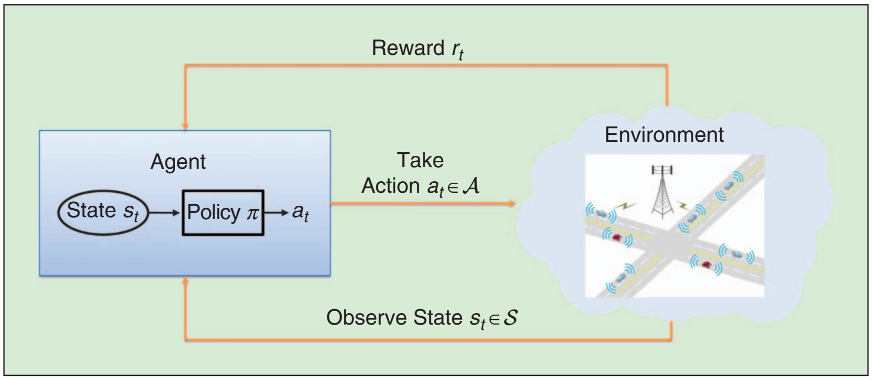

FIGURE 4 The deep reinforcement learning for V2V communications.

We assume the vehicle-to-infrastructure (V2I) link has been allocated orthogonal resources beforehand, and the main goal of the proposed distributed channel and power allocation is to ensure the latency constraints for each V2V link and minimize interference to V2I links. The structure of reinforcement learning for V2V communications is shown in Figure 4, where an agent, corresponding to a V2V link, interacts with the environment. In this scenario, the environment is considered to be everything outside the V2V link. It should be noted that the behaviors of other V2V links cannot be controlled in the decentralized settings. As a result, their actions, such as selected spectrum and transmission power, can only be treated as part of the environment.

As in Figure 4, at time t, each V2V link, as an agent, observes a state, s_t, from the state space, $\mathcal{S}$ and accordingly takes an action, a_t, selected from the action space, $\mathcal{A}$, which amounts to selecting the subband and transmission power based on the policy, π. The decision policy, π, is determined by a Q function, $Q(s_t, a_t, \theta)$, where θ is the parameter of the Q function and can be obtained by deep learning. Following the action, the state of the environment transitions to a new state, s_{t+1}, and the agent receives a reward, r_t, determined by the capacity of the V2I links and the corresponding V2V latency. In our system, the state observed by each V2V link for characterizing the environment consists of several parts: the instantaneous channel information of the corresponding V2V link, g_t; the previous interference to the link, I_{t-1}; the channel information of the V2I link, i.e., from the V2V transmitter to the base station, h_t; the selection of subbands of neighbors in the previous time slot, B_{t-1}; the remaining load for the vehicles to transmit, L_t; and the remaining time to meet the latency constraints, U_t. Hence, the state can be expressed as $s_t = [g_t, I_{t-1}, h_t, B_{t-1}, L_t, U_t]$. The training and testing data are generated from an environment simulator based on 3GPP channel models. In the training stage, we follow the deep Q learning with experience replay, where the generated data are saved in a storage called *memory*. The minibatch data used for updating the Q network are sampled from the memory. In this way, the temporal correlation of data can be suppressed. The policy used in each V2V link for selecting spectrum and power is random at the beginning and gradually improved with the updated Q networks.

Figure 5 compares the proposed method with a random resource allocation method, where the agent randomly chooses a subband for transmission at each time. From the figure, the previous reinforcement learning-based method has a larger probability for V2V links to satisfy the latency constraint because it can dynamically adjust the power and subband for transmission so that the links that are likely to violate the latency constraint have more resources.

Open Issues

Recent hype around machine learning seems to suggest it is a panacea to most conventionally challenging problems, especially in view of the significant advances made by deep learning. However, naively applying existing machine-learning methods to vehicular networks is expected to be insufficient due to their many distinguishing characteristics. How to adapt existing learning methods or develop V2X-specific learning algorithms to better handle such characteristics remains a challenging task. Therefore, in this section, we identify several research topics for further investigation.

Learning Dynamics of Vehicular Networks

Vehicular networks exhibit strong dynamics in many facets, e.g., wireless propagation channels, network

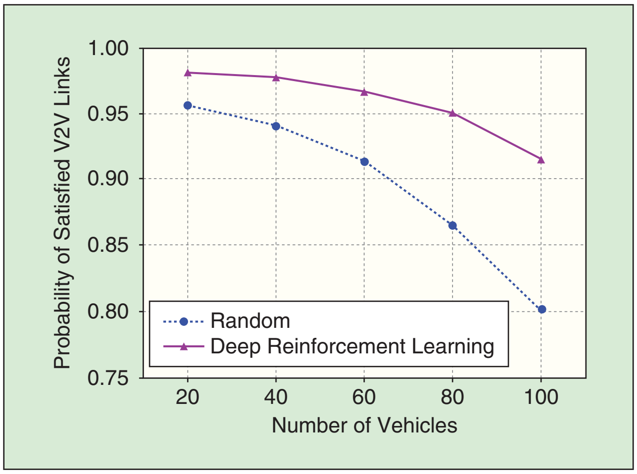

FIGURE 5 The success probabilities versus the number of vehicles.

topologies, and traffic dynamics. How to efficiently learn and robustly predict such dynamics based on historical data generated from multiple onboard sensors or previous transmission is still an open issue. Traditionally, some Bayesian models, such as hidden Markov models, can be used for characterizing the temporal relationship and predicting states in the next time slot. New sophisticated models powered by deep neural networks, such as RNN and LSTM, can improve the prediction by exploiting the long-range dependency.

One potential application is to predict wireless channels using deep neural networks based on the received signal and historical data. Due to the high-mobility of users and advanced techniques being adopted, such as massive multiple-input, multiple-output and millimeter-wave communications, new challenges arise for estimating high-dimensional fast-varying wireless channels. Deep learning has shown strong abilities to efficiently distill high-dimensional data by exploiting such properties as sparsity. It is unclear whether deep neural networks can assist or even replace the existing channel estimation module that requires transmitting frequent pilot symbols to track channel variations. Another potential application is to predict the vehicular trajectory, which could be further used for traffic dynamics prediction. The latent factors that affect the trajectories, such as drivers' intention, traffic situations, and road structures, may be implicitly learned from the historical data based on deep neural networks. More research efforts are thus needed to develop a better understanding in this area.

Method Complexity

Conventional machine-learning techniques require much work on designing feature representation, whereas deep learning provides an effective way to learn features from raw data. With deep models, information can be extracted more efficiently than the traditional methods, and experimental results have confirmed that the deep hierarchical structure is necessary for better understanding of data. Key technologies have since been devised to get deeper models with stronger representation abilities. With high-performance computing facilities, such as graphics processing units, deeper networks can be trained with massive amounts of data through advanced training techniques, such as batch norm and residual networks [5]. However, in vehicular networks, onboard computation resources are limited, and the low end-to-end latency requirement restricts heavy use of cloud servers housed remotely. It is therefore important to develop special treatments for vehicular networks, such as model reduction or compression, to relieve the resource limitation without compromising performance.

Distributed Representation

In vehicular networks, data are naturally generated and stored across different units in the network, e.g., vehicles, roadside units, and remote clouds. This brings challenges to the applicability of most existing machine-learning algorithms that have been developed under the assumption that data are centrally controlled and easily accessible. As a result, distributed learning methods are desired in vehicular networks that act on partially observed data and have the ability to exploit information obtained from other entities in the network. Furthermore, additional overheads incurred by the coordination and sharing of information among various units in vehicular networks for distributed learning shall be properly accounted for to make the system work effectively.

Conclusions

This article provides an overview of applying machine learning to address challenges in emerging vehicular networks. We briefly introduced the basics of machine learning, including major categories and representative algorithms. We provided some preliminary examples in applying machine learning in vehicular networks to facilitate data-driven decision making and discussed intelligent wireless resource management using reinforcement learning in detail. We also highlighted some open issues for further research.

Acknowledgment

This work was supported in part by a research gift from Intel Corporation and the National Science Foundation under Grants 1443894 and 1731017.

Author Information

Hao Ye (yehao@gatech.edu) received his B.E. degree in information engineering from Southeast University, Nanjing, China, in 2012, and his M.E. degree in electrical engineering from Tsinghua University, Beijing, China, in 2015. He is currently working toward his Ph.D. degree in electrical engineering at the Georgia Institute of Technology,

Atlanta. His research interests include machine learning, signal detection, vehicular communications, and radio resource management.

Le Liang (lliang@gatech.edu) received his B.E. degree in information engineering from Southeast University, Nanjing, China, in 2012, and his M.A.Sc. degree in electrical engineering from the University of Victoria, British Columbia, Canada, in 2015. He is currently working toward his Ph.D. degree in electrical engineering at the Georgia Institute of Technology, Atlanta. His research interests include vehicular communications, radio resource management, and multiple-input, multiple-output systems.

Geoffrey Ye Li (liye@ece.gatech.edu) is currently a professor with the Georgia Institute of Technology (Georgia Tech), Atlanta. His research interests include signal processing and machine learning for wireless communications. In these areas, he has published more than 400 articles with more than 30,000 citations and has been listed as a highly cited researcher by Thomson Reuters. He has won the IEEE Communications Society Stephen O. Rice Prize Paper Award and Award for Advances in Communication, the IEEE Vehicular Technology Society James Evans Avant Garde Award and Jack Neubauer Memorial Award, the IEEE Signal Processing Society Donald G. Fink Overview Paper Award, and the Distinguished Electrical and Computer Engineering Faculty Achievement Award from Georgia Tech. He is a Fellow of the IEEE.

JoonBeom Kim (joonbeom.kim@intel.com) received his Ph.D. degree in electrical engineering from the Georgia Institute of Technology, Atlanta, in 2006. He has been a research scientist at Intel Corporation, Santa Clara, California, working on fifth-generation wireless system design since 2014. His research interests include several aspects of signal processing and wireless communication: channel estimation; synchronization; iterative processing; channel coding; orthogonal frequency-division multiplexing; multiple-input, multiple-output processing; physical-layer security; and machine learning for wireless communications.

Lu Lu (lu2.lu@intel.com) received her B.S.E. degree and M.S.E. degree from the University of Electronic Science and Technology of China, Chengdu, in 2007 and 2010, respectively; her licentiate degree from the Royal Institute of Technology, Stockholm, Sweden, in 2011; and her Ph.D. degree in electrical engineering from the Georgia Institute of Technology, Atlanta, in 2015. She is currently a research scientist at Intel Labs, Hillsboro, Oregon. Her research interests include future wireless technology, standards, and signal processing.

May Wu (may.wu@intel.com) received her Ph.D. degree in electrical engineering. She is currently a senior principal engineer and director of wireless systems and architecture for Intel Labs communications group, Santa Clara, California. Her research interests include system-on-a-chip architecture, large-scale integrated circuit research and design, microprocessors, and embedded systems.

References

[1] L. Liang, H. Peng, G. Y. Li, and X. Shen, “Vehicular communications: A physical layer perspective,” *IEEE Trans. Veh. Technol.*, vol. 66, no. 12, pp. 10,647–10,659, Dec. 2017.
[2] H. Peng, L. Liang, X. Shen, and G. Y. Li. (2017). Vehicular communications: A network layer perspective. arXiv. [Online]. Available: https://arxiv.org/abs/1707.09972
[3] E. Alpaydin, *Introduction to Machine Learning*. Cambridge, MA: MIT Press, 2014.
[4] J. Schmidhuber, “Deep learning in neural networks: An overview,” *Neural Netw.*, vol. 61, pp. 85–117, Jan. 2015.
[5] C. Jiang, H. Zhang, Y. Ren, Z. Han, K.-C. Chen, and L. Hanzo, “Machine learning paradigms for next-generation wireless networks,” *IEEE Wireless Commun.*, vol. 24, no. 2, pp. 98–105, Apr. 2017.
[6] C. Ide, F. Hadiji, L. Habel, A. Molina, T. Zaksek, M. Schreckenberg, K. Kersting, and C. Wietfeld, “LTE connectivity and vehicular traffic prediction based on machine learning approaches,” in *Proc. IEEE Vehicular Technology Conf. (VTC Fall)*, 2015, pp. 1–5.
[7] Y. Lv, Y. Duan, W. Kang, Z. Li, and F.-Y. Wang, “Traffic flow prediction with big data: A deep learning approach,” *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 2, pp. 865–873, Apr. 2015.
[8] C. Wu, T. Yoshinaga, Y. Ji, T. Murase, and Y. Zhang, “A reinforcement learning-based data storage scheme for vehicular ad hoc networks,” *IEEE Trans. Veh. Technol.*, vol. 66, no. 7, pp. 6336–6348, July 2017.
[9] N. Taherkhani and S. Pierre, “Centralized and localized data congestion control strategy for vehicular ad hoc networks using a machine learning clustering algorithm,” *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 11, pp. 3275–3285, Nov. 2016.
[10] Z. Li, C. Wang, and C.-J. Jiang, “User association for load balancing in vehicular networks: An online reinforcement learning approach,” *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 8, pp. 2217–2228, Aug. 2017.
[11] Y. Xu, L. Li, B.-H. Soong, and C. Li, “Fuzzy Q-learning based vertical handoff control for vehicular heterogeneous wireless network,” in *Proc. IEEE Int. Conf. Communications (ICC)*, 2014, pp. 5653–5658.
[12] M. A. Salahuddin, A. Al-Fuqaha, and M. Guizani, “Reinforcement learning for resource provisioning in the vehicular cloud,” *IEEE Wireless Commun.*, vol. 23, no. 4, pp. 128–135, June 2016.
[13] Q. Zheng, K. Zheng, H. Zhang, and V. C. Leung, “Delay-optimal virtualized radio resource scheduling in software-defined vehicular networks via stochastic learning,” *IEEE Trans. Veh. Technol.*, vol. 65, no. 10, pp. 7857–7867, Oct. 2016.
[14] T. He, N. Zhao, and H. Yin, “Integrated networking, caching and computing for connected vehicles: A deep reinforcement learning approach,” *IEEE Trans. Veh. Technol*, vol. 67, no. 1, pp. 44–55, Oct. 2017.
[15] H. Ye, G. Y. Li, and B.-H. Juang. (2017). Deep reinforcement learning for resource allocation in V2V communications. arXiv. [Online]. Available: https://arxiv.org/abs/1711.00968

VT