Wireless for Machine Learning: A Survey

Other titles in Foundations and Trends® in Signal Processing

Bilevel Methods for Image Reconstruction
Caroline Crockett and Jeffrey A. Fessler
ISBN: 978-1-63828-002-6

Operating Characteristics for Classical and Quantum Binary Hypothesis Testing
Catherine A. Medlock and Alan V. Oppenheim
ISBN: 978-1-68083-882-4

Foundations of User-Centric Cell-Free Massive MIMO
Özlem Tugfe Demir, Emil Björnson and Luca Sanguinetti
ISBN: 978-1-68083-790-2

Data-Driven Multi-Microphone Speaker Localization on Manifolds
Bracha Laufer-Goldshtein, Ronen Talmon and Sharon Gannot
ISBN: 978-1-68083-736-0

Recent Advances in Clock Synchronization for Packet-Switched Networks
Anantha K. Karthik and Rick S. Blum
ISBN: 978-1-68083-726-1

Wireless for Machine Learning: A Survey

Henrik Hellström
KTH Royal Institute of Technology
hhells@kth.se

José Mairton B. da Silva Jr.
KTH Royal Institute of Technology

Mohammad Mohammadi Amiri
Massachusetts Institute of Technology

Mingzhe Chen
Princeton University

Viktoria Fodor
KTH Royal Institute of Technology

H. Vincent Poor
Princeton University

Carlo Fischione
KTH Royal Institute of Technology

now
the essence of knowledge
Boston — Delft

Foundations and Trends® in Signal Processing

Published, sold and distributed by:
now Publishers Inc.
PO Box 1024
Hanover, MA 02339
United States
Tel. +1-781-985-4510
www.nowpublishers.com
sales@nowpublishers.com

Outside North America:
now Publishers Inc.
PO Box 179
2600 AD Delft
The Netherlands
Tel. +31-6-51115274

The preferred citation for this publication is

H. Hellström *et al.*. *Wireless for Machine Learning: A Survey*. Foundations and Trends® in Signal Processing, vol. 15, no. 4, pp. 290–399, 2022.

ISBN: 978-1-63828-007-1

Foundations and Trends® in Signal Processing
Volume 15, Issue 4, 2022

Editorial Scope

Topics

Foundations and Trends® in Signal Processing publishes survey and tutorial articles in the following topics:

- Adaptive signal processing
- Audio signal processing
- Biological and biomedical signal processing
- Complexity in signal processing
- Digital signal processing
- Distributed and network signal processing
- Image and video processing
- Linear and nonlinear filtering
- Multidimensional signal processing
- Multimodal signal processing
- Multirate signal processing
- Multiresolution signal processing
- Nonlinear signal processing
- Randomized algorithms in signal processing
- Sensor and multiple source signal processing, source separation
- Signal decompositions, subband and transform methods, sparse representations
- Signal processing for communications
- Signal processing for security and forensic analysis, biometric signal processing
- Signal quantization, sampling, analog-to-digital conversion, coding and compression
- Signal reconstruction, digital-to-analog conversion, enhancement, decoding and inverse problems
- Speech/audio/image/video compression
- Speech and spoken language processing
- Statistical/machine learning
- Statistical signal processing
  - Classification and detection
  - Estimation and regression
  - Tree-structured methods

Information for Librarians

Foundations and Trends® in Signal Processing, 2022, Volume 15, 4 issues. ISSN paper version 1932-8346. ISSN online version 1932-8354. Also available as a combined paper and online subscription.

Contents

Wireless for Machine Learning: A Survey

Henrik Hellström[1], José Mairton Barros da Silva Jr.[1], Mohammad Mohammadi Amiri[2], Mingzhe Chen[3], Viktoria Fodor[1], H. Vincent Poor[3] and Carlo Fischione[1]

[1]*School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology, Sweden; hhells@kth.se*
[2]*MIT Media Laboratory, Massachusetts Institute of Technology, USA*
[3]*Department of Electrical and Computer Engineering, Princeton University, USA*

ABSTRACT

As data generation increasingly takes place on devices without a wired connection, Machine Learning (ML) related traffic will be ubiquitous in wireless networks. Many studies have shown that traditional wireless protocols are highly inefficient or unsustainable to support ML, which creates the need for new wireless communication methods. In this monograph, we give a comprehensive review of the state-of-the-art wireless methods that are specifically designed to support ML services over distributed datasets. Currently, there are two clear themes within the literature, analog over-the-air computation and digital radio resource management optimized for ML. This survey gives an introduction to these methods, reviews the most important works, highlights open problems, and discusses application scenarios.

Henrik Hellström, José Mairton Barros da Silva Jr., Mohammad Mohammadi Amiri, Mingzhe Chen, Viktoria Fodor, H. Vincent Poor and Carlo Fischione (2022), "Wireless for Machine Learning: A Survey", Foundations and Trends® in Signal Processing: Vol. 15, No. 4, pp 290–399. DOI: 10.1561/2000000114.

1

Introduction

With the increasing popularity of mobile devices and the continuous growth of Internet of Things (IoT), we are having increasing access to vast amounts of distributed data. According to a recent report from Ericsson, the global number of connected IoT devices will rise to 4.1 billion by 2024 [49], which is four times the 1 billion observed in 2019. Simultaneously, breakthroughs in Machine Learning (ML) are allowing us to analyze the data of edge devices so as to solve a wide range of complex problems, such as image recognition [66], language processing [39], and predictive modeling [23]. However, since ML was originally conceived in centralized settings where all data must be aggregated at a common location, the application of ML on distributed datasets over wireless networks is generating new challenges for the wireless networks, namely:

- **Privacy:** Many ML applications require the use of privacy-sensitive data. In these cases, it is either desirable or necessary that the training dataset cannot be inferred by eavesdropping upon the ML updates being transferred wirelessly [150];
- **Security:** When an ML model is trained distributively, a bad actor can corrupt the final model by transmitting malicious model updates [159]. Wireless protocol design should inhibit an attacker's ability to do so;

- **Communication and Energy Efficiency:** Distributed ML (DML) requires the communication of high-dimensional model updates for hundreds or thousands of iterations before the model has converged. This communication of updates generally forms the performance bottleneck of the training process, imposing the risk of excessively draining the batteries of training devices and overwhelming the capacity of the wireless network [144].

To address these challenges, a new approach toward communication protocol design has emerged [198]. This new approach considers the design of new wireless methods for carrying data needed for the ML tasks. Unlike traditional wireless protocol design, the objective of Wireless for ML is not to deliver bits as efficiently as possible, but to distill the intelligence carried within the data. The traditional communication protocols that are designed to maximize data rate and minimize bit errors have been shown to be greatly inefficient for carrying ML related data [9], [35], [100], [118], [200]. Instead, Wireless for ML offers new methods that are better aligned with the ML objective and invites us to rethink how wireless communication protocols are designed. Among the novel methods that have been proposed, two major themes arise, namely analog over-the-air computation (AirComp) and radio resource management (RRM) optimized for ML. In AirComp, the long-standing doctrine of interference avoidance is questioned and novel interference-promoting protocols are proposed while in RRM for ML, the new objectives lead to solutions that are fundamentally different from what is used today.

The idea of wireless protocols customized for ML, although not yet available in the current cellular wireless standards, is compatible with the current standard specifications. The new cellular standard 5G has introduced the concept of network slicing to improve flexibility and scalability [130]. Network slicing allows independent sets of network protocols to run on common physical infrastructure, to support services with conflicting requirements. As an example, video streaming requires high data rates and accepts high latency, while critical IoT usually requires low latency and high reliability while accepting low data rates. Prior to the emergence of 5G, these services could not be supported using the same protocols, but with network slicing, they can be implemented on the same physical infrastructure [15]. Going beyond 5G, the demand for ML services is projected to grow significantly and discussions have begun on a dedicated network slice for ML in future-generation cellular networks such as beyond-5G and 6G [60], [131], [151], [191]. Given this possibility, the investigation of Wireless for ML becomes relevant not only for local-area networks but also for large-scale cellular networks.

1.1 Related Work

Although the general intersection of ML and wireless communications is currently a prolific field of research that has already generated multiple surveys, there are fewer works reviewing Wireless for ML. The current surveys can roughly be classified into three categories: *ML for Wireless Communications*, *Wireless for ML*, and *Communication-Efficient DML*. We list a set of representative surveys in Table 1.1. A brief description of the three areas follows.

1. **Wireless for ML** uses wireless communication protocols as a method to enable or significantly improve ML training over wireless networks. Unlike in traditional wireless communication, the communication system is not oblivious to the meaning that the bits convey. Instead, Wireless for ML is a task-oriented communication philosophy, where the goal of the communication system is to distill the intelligence carried within the data.

2. **Communication-efficient DML** has the same goal as Wireless for ML but uses different methods. Instead of customizing the wireless protocols, advancements are made by modifying or redesigning the ML algorithm. The results of these works are agnostic to the communication protocol so that they can be applied regardless of the specific technologies used to transmit data.

3. **ML for wireless** uses ML as a method to design wireless communication protocols or other elements for general communication services. Therefore, its goal is the same as in traditional wireless communications, i.e., efficient and reliable transfer of arbitrary data. The communication system should support a wide variety of services and is therefore deliberately oblivious to the semantics of transmitted bits.

Table 1.1: Surveys written within the intersection of ML and communications. The topics of ML for Communications and Communication-efficient DML have been covered in many surveys, unlike Wireless for ML. At most, Wireless for ML has been covered briefly in conjunction with Communication-efficient DML.

Year	Journal	Ref.	Research Area from Figure 1.1
2017	IEEE Communication Surveys and Tutorials	[109]	3
2018	Proceedings of the IEEE	[120]	2
2019	Proceedings of the IEEE	[194]	2
2020	IEEE Communication Surveys and Tutorials	[73]	3
2020	IEEE Communication Surveys and Tutorials	[162]	3
2020	IEEE Internet of Things Journal	[40]	Mostly 2 with some 1
2020	IEEE Communication Surveys and Tutorials	[164]	2
2020	IEEE Internet of Things Journal	[3]	2
2020	IEEE Communication Surveys and Tutorials	[178]	Mostly 2 with some 1
2021	IEEE Internet of Things Journal	[74]	2
2021	Elsevier High-Confidence Computing	[170]	2
2021	arXiv	[54]	Mostly 1 with some 2
This survey			1

In addition to the three categories above, their intersections can be considered as areas of their own, illustrated in Figure 1.1. The intersection of Wireless for ML and Communication-efficient DML considers the co-design of the ML algorithm and the wireless protocol. With such an approach, researchers attempt to reach some global optimality, which is lost when the two problems are treated in isolation. Additionally, one can consider the intersection between Wireless for ML and ML for Wireless, where ML would be used as a tool to design a wireless protocol with the goal of supporting distributed ML services. However, as far as we are aware, no works have been published in this direction. In this survey, we consider all works within Wireless for ML, including its intersections, symbolized by the green crescent in Figure 1.1.

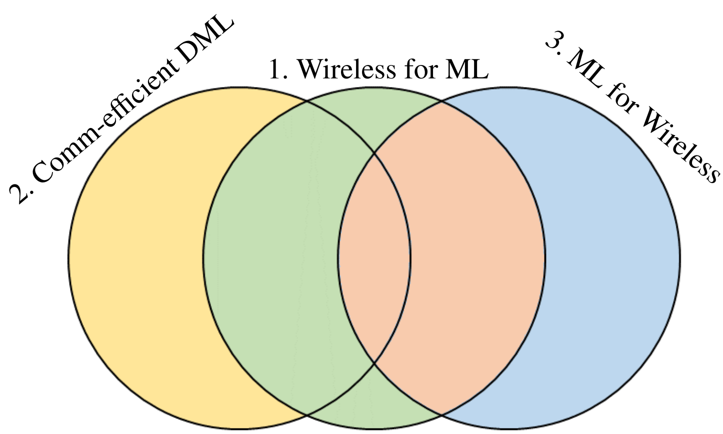

Figure 1.1: Illustration of the relationship between Wireless for ML and related fields. The first circle correponds to Communication-efficient DML, the second to Wireless for ML, and the third to ML for Wireless. The blue area corresponds to pure ML for Wireless, which is a very prolific field of research that has already generated a large number of review articles. Likewise, the yellow area corresponds to pure Communication-efficient DML which is also a well-covered area. In this survey, we focus on the green moon, i.e., pure Wireless for ML and its intersection with Communication-efficient DML. As far as we are aware, there are no published works in the red area.

Some of the papers in Table 1.1 discuss Wireless for ML, but the treatments there are not extensive since that is not the main purpose of these papers. The closest match to our survey is [54]. However, despite describing some works within Wireless for ML, the paper is not a comprehensive survey of the field, instead its purpose is to introduce a new framework to describe Federated

Learning. We believe that due to this gap, there is currently no one-stop survey that offers an overview of the Wireless for ML literature, which motivates us to write this survey with the following contributions:

- We provide an introduction to important concepts necessary to understand the field as a whole, such as DML, over-the-air computation, and the distinction between generic wireless communication protocols and Wireless for ML;
- We describe the most important works of the field in a concise way to offer a thorough overview of the state-of-the-art, both for analog over-the-air computation and digital communications;
- We discuss several important open problems and future research directions within Wireless for ML;
- We describe a number of application areas where Wireless for ML can provide a benefit to society, such as vehicular communications and virtual reality, and describe the challenges associated with those applications.

1.2 Notation and Organization

All the contributions that we survey are essentially concerned with the solution to a basic problem, namely the training of a classifier over a wireless communication network constrained by the natural characteristics of the wireless channel. Throughout this survey, we assume a centralized architecture where there is a central controller or parameter server (PS) able to make decisions such as user selection, bandwidth allocation, and aggregation frequency control. Such an architecture is representative of most of the wireless networks used today, from large scale mobile to personal area networks. The communication channel is wireless and is thus subject to fading, additive noise, and bandwidth restrictions. The training dataset is always carried by user devices and the training algorithms will always be chosen to minimize a loss based on the global dataset. Unless specified otherwise, the network consists of one PS, i.e., the base station (BS) or the access point (AP), and K user devices, e.g., IoT devices, user equipments (UEs), or other wireless devices. Each device (say the k^{th}) carries a subset $\mathcal{D}_k$ of the global dataset $\mathcal{D}$ and the PS carries no data. The global dataset consists of N training samples and corresponds to the union of data available at all the user devices. For communication, the uplink h_k and downlink g_k channel coefficients corresponding to the k^{th} UE are of particular importance. Figure 1.2 illustrates the setup, a full list of notation is given in Table 1.2, and relevant abbreviations are given in Table 1.3.

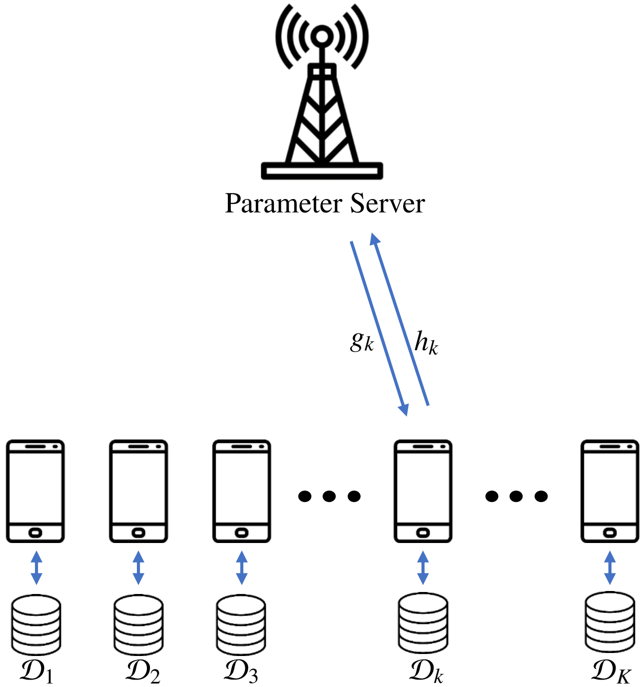

Figure 1.2: Illustration of the PS and wireless network setup used throughout this survey. Current wireless communication protocols substantially hinder or completely block distributed training over this setup. The Wireless for ML paradigm is an approach to tackle such hindrances and blockages.

The rest of this survey is organized as follows: Section 2 provides a primer on DML and in particular Federated Learning (FL). In Sections 3 and 4, we survey the Wireless for ML literature for over-the-air computation and digital communication, respectively. In Section 5, we discuss the open problems in Wireless for ML within both analog over-the-air computation and digital communications. Then, in Section 6, we discuss applications supported by Wireless for ML. Finally, we give some concluding remarks in Section 7.

Table 1.2: Reference list of commonly used variables in this survey. Ordered alphabetically and by case.

Variable	Interpretation
B	Bandwidth available to the learning system
$\mathcal{D}_k$	Dataset carried by device k
E	Number of epochs
K	Number of user devices
M	Number of antennas at the parameter server
N	Number of data samples in the global dataset
N_k	Number of data samples stored at device k
$\mathcal{S}^t$	Set of selected devices at iteration t
T_{round}	Time for federated learning communication round
β	Learning rate
η	Post-transmission scalar
$\nabla f(\mathbf{w})$	Gradient of function f evaluated at $\mathbf{w}$
b_k	Ratio of total bandwidth allocated to device k
d	Number of model parameters in $\mathbf{w}$
$f(\mathbf{w})$	Empirical risk function of the global model $\mathbf{w}$
g_k	CSI in downlink direction from server to device k
h_k	CSI in uplink direction from device k to server
$l(\mathbf{w})$	Loss function for parameter $\mathbf{w}$
p_k	Uplink power allocated to device k
v	Additive white Gaussian noise
$\mathbf{w}^t$	Global model parameters at iteration t
$\mathbf{w}_k^t$	Local model parameters for device k at iteration t
$\mathbf{x}$	Input or feature of data sample
$\mathbf{y}$	Output or label of data sample

Table 1.3: Reference list of most abbreviations used in this survey.

Acronym	**Phrase**
ADMM	Alternating Direction Method of Multipliers
AirComp	Over-the-air Computation
BAA	Broadband Analog Aggregation
BPSK	Binary Phase-Shift Keying
BS	Base Station
CML	Centralized Machine Learning
CoCoA	Comm-efficient distributed dual Coordinate Ascent
CoMAC	Computation over Multiple-Access Channels
CSI	Channel State Information
DML	Distributed Machine Learning
DP	Differential Privacy
DSGD	Distributed Stochastic Gradient Descent
ESN	Echo State Network
FD	Federated Distillation
FedAvg	Federated Averaging
FL	Federated Learning
IID	Independent and Identically Distributed
IRS	Intelligent Reflective Surface
IoT	Internet of Things
LTE	Long Term Evolution
MIMO	Multiple Input Multiple Output
ML	Machine Learning
MSE	Mean Square Error
OFDMA	Orthogonal Frequency Division Multiple Access
PS	Parameter Server
RRM	Radio Resource Management
SGD	Stochastic Gradient Descent
SISO	Single Input Single Output
SNR	Signal to Noise Ratio
QoE	Quality of Experience
UAV	Unmanned Aerial Vehicle
VR	Virtual Reality
ZF	Zero-Forcing

References

[1] O. Abari, H. Rahul, and D. Katabi, "Over-the-Air Function Computation in Sensor Networks," *arXiv abs/1612.02307*, 2016.

[2] O. Abari, H. Rahul, D. Katabi, and M. Pant, "Airshare: Distributed Coherent Transmission Made Seamless," in *Proceedings of the 2015 IEEE Conference on Computer Communications (INFOCOM)*, IEEE, pp. 1742–1750, 2015.

[3] S. Abdulrahman, H. Tout, H. Ould-Slimane, A. Mourad, C. Talhi, and M. Guizani, "A Survey on Federated Learning: The Journey from Centralized to Distributed On-Site Learning and Beyond," *IEEE Internet of Things Journal*, vol. 8, no. 7, 2020, pp. 5476–5497.

[4] J.-H. Ahn, O. Simeone, and J. Kang, "Wireless Federated Distillation for Distributed Edge Learning with Heterogeneous Data," in *Proceedings of the 2019 IEEE 30th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, IEEE, pp. 1–6, 2019.

[5] J.-H. Ahn, O. Simeone, and J. Kang, "Cooperative Learning via Federated Distillation over Fading Channels," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 8856–8860, Barcelona, Spain, 2020.

[6] M. M. Amiri, T. M. Duman, D. Gündüz, S. R. Kulkarni, and H. V. Poor, "Blind Federated Edge Learning," *IEEE Transactions on Wireless Communications*, 2021.

[7] M. M. Amiri, D. Gündüz, S. R. Kulkarni, and H. V. Poor, "Update Aware Device Scheduling for Federated Learning at the Wireless Edge," in *Proceedings of the IEEE International Symposium on Information Theory*, pp. 2598–2603, Los Angeles, CA, USA, 2020.

[8] M. M. Amiri, D. Gündüz, S. R. Kulkarni, and H. V. Poor, "Convergence of Update Aware Device Scheduling for Federated Learning at the Wireless Edge," *IEEE Transactions on Wireless Communications*, vol. 20, no. 6, 2021, pp. 3643–3658.

[9] M. M. Amiri and D. Gunduz, "Machine Learning at the Wireless Edge: Distributed Stochastic Gradient Descent Over-the-Air," *IEEE Transactions on Signal Processing*, vol. 68, 2020, pp. 2155–2169.

[10] M. M. Amiri and D. Gündüz, "Over-the-Air Machine Learning at the Wireless Edge," in *Proceedings of the IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, IEEE, pp. 1–5, 2019.

[11] M. M. Amiri and D. Gündüz, "Federated Learning over Wireless Fading Channels," *IEEE Transactions on Wireless Communications*, vol. 19, no. 5, 2020, pp. 3546–3557.

[12] M. M. Amiri, D. Gündüz, S. R. Kulkarni, and H. V. Poor, "Convergence of Federated Learning over a Noisy Downlink," *IEEE Transactions on Wireless Communications*, 2021.

[13] E. Bagdasaryan, A. Veit, Y. Hua, D. Estrin, and V. Shmatikov, "How to Backdoor Federated Learning," in *Proceedings of the International Conference on Artificial Intelligence and Statistics*, PMLR, pp. 2938–2948, 2020.

[14] T. Ben-Nun and T. Hoefler, "Demystifying Parallel and Distributed Deep Learning: An In-Depth Concurrency Analysis," *ACM Computing Surveys*, vol. 52, no. 4, 2019.

[15] M. Bennis, M. Debbah, and H. V. Poor, "Ultrareliable and Low-Latency Wireless Communication: Tail, Risk, and Scale," *Proceedings of the IEEE*, vol. 106, no. 10, 2018, pp. 1834–1853.

[16] J. Bernstein, Y.-X. Wang, K. Azizzadenesheli, and A. Anandkumar, "signSGD: Compressed Optimisation for Non-Convex Problems," in *Proceedings of the International Conference on Machine Learning*, PMLR, pp. 560–569, 2018.

[17] T. Bertin-Mahieux, D. P. Ellis, B. Whitman, and P. Lamere, "The Million Song Dataset," in *Proceedings of the 11th International Conference on Music Information Retrieval (ISMIR)*, 2011.

[18] A. N. Bhagoji, S. Chakraborty, P. Mittal, and S. Calo, "Analyzing Federated Learning Through an Adversarial Lens," in *Proceedings of the International Conference on Machine Learning*, PMLR, pp. 634–643, 2019.

[19] P. Blanchard, E. M. El Mhamdi, R. Guerraoui, and J. Stainer, "Machine Learning with Adversaries: Byzantine Tolerant Gradient Descent," *Advances in Neural Information Processing Systems*, vol. 30, 2017.

[20] P. Bory, "Deep New: The Shifting Narratives of Artificial intelligence from Deep Blue to AlphaGo," *Convergence: The International Journal of Research into New Media Technologies*, vol. 25, no. 4, 2019, pp. 627–642.

[21] L. Bottou, F. E. Curtis, and J. Nocedal, "Optimization Methods for Large-Scale Machine Learning," *SIAM Review*, vol. 60, no. 2, 2018, pp. 223–311.

[22] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, *et al.*, "Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers," *Foundations and Trends® in Machine Learning*, vol. 3, no. 1, 2011, pp. 1–122.

[23] P. Branco, L. Torgo, and R. P. Ribeiro, "A Survey of Predictive Modeling on Imbalanced Domains," *ACM Computing Surveys (CSUR)*, vol. 49, no. 2, 2016, pp. 1–50.

[24] B. Brik, A. Ksentini, and M. Bouaziz, "Federated Learning for UAVs-enabled Wireless Networks: Use Cases, Challenges, and Open Problems," *IEEE Access*, vol. 8, 2020, pp. 53 841–53 849.

[25] R. W. Broadley, J. Klenk, S. B. Thies, L. P. Kenney, and M. H. Granat, "Methods for the Real-World Evaluation of Fall Detection Technology: A Scoping Review," *Sensors*, vol. 18, no. 7, 2018, p. 2060.

[26] S. Caldas, J. Konecny, H. B. McMahan, and A. Talwalkar, "Expanding the Reach of Federated Learning by Reducing Client Resource Requirements," *arXiv abs/1812.07210*, 2019.

[27] X. Cao, G. Zhu, J. Xu, and S. Cui, "Transmission Power Control for Over-the-Air Federated Averaging at Network Edge," *IEEE Journal on Selected Areas in Communications*, 2022, pp. 1571–1586.

[28] X. Cao, G. Zhu, J. Xu, and S. Cui, “Optimized Power Control for Over-the-Air Federated Edge Learning,” in *Proceedings of the 2021 IEEE International Conference on Communications*, IEEE, pp. 1–6.

[29] X. Cao, G. Zhu, J. Xu, and K. Huang, “Optimized Power Control for Over-the-Air Computation in Fading Channels,” *IEEE Transactions on Wireless Communications*, vol. 19, no. 11, 2020, pp. 7498–7513.

[30] N. Carlini, C. Liu, Ú. Erlingsson, J. Kos, and D. Song, “The Secret Sharer: Evaluating and Testing Unintended Memorization in Neural Networks,” in *Proceedings of the 28th {USENIX} Security Symposium ({USENIX} Security 19)*, pp. 267–284, 2019.

[31] F. Chen, M. Luo, Z. Dong, Z. Li, and X. He, “Federated Meta-Learning with Fast Convergence and Efficient Communication,” *abs/1802.07876*, 2018.

[32] M. Chen, H. V. Poor, W. Saad, and S. Cui, “Convergence Time Optimization for Federated Learning Over Wireless Networks,” *IEEE Transactions on Wireless Communications*, vol. 20, no. 4, 2021, pp. 2457–2471.

[33] M. Chen, O. Semiari, W. Saad, X. Liu, and C. Yin, “Federated Echo State Learning for Minimizing Breaks in Presence in Wireless Virtual Reality Networks,” *IEEE Transactions on Wireless Communications*, vol. 19, no. 1, 2020, pp. 177–191.

[34] M. Chen, N. Shlezinger, H. V. Poor, Y. C. Eldar, and S. Cui, “Communication-efficient Federated Learning,” *Proceedings of the National Academy of Sciences*, vol. 118, no. 17, 2021.

[35] M. Chen, Z. Yang, W. Saad, C. Yin, H. V. Poor, and S. Cui, “A Joint Learning and Communications Framework for Federated Learning over Wireless Networks,” *IEEE Transactions on Wireless Communications*, vol. 20, no. 1, 2020, pp. 269–283.

[36] W. Chen and H. V. Poor, *Edge Caching for Mobile Networks*. London, UK: IET Press, 2021.

[37] Y. Cheng, D. Wang, P. Zhou, and T. Zhang, “A Survey of Model Compression and Acceleration for Deep Neural Networks,” *arXiv abs/1710.09282*, 2017.

[38] B. Clerckx, K. Huang, L. R. Varshney, S. Ulukus, and M.-S. Alouini, "Wireless Power Transfer for Future Networks: Signal Processing, Machine Learning, Computing, and Sensing," *arXiv abs/2101.04810*, 2021.

[39] R. Collobert, J. Weston, L. Bottou, M. Karlen, K. Kavukcuoglu, and P. Kuksa, "Natural Language Processing (almost) from Scratch," *Journal of Machine Learning Research*, vol. 12, 2011, pp. 2493–2537.

[40] S. Deng, H. Zhao, W. Fang, J. Yin, S. Dustdar, and A. Y. Zo5a, "Edge Intelligence: the Confluence of Edge Computing and Artificial Intelligence," *IEEE Internet of Things Journal*, 2020.

[41] J. Ding, R. Calderbank, and V. Tarokh, "Gradient Information for Representation and Modeling," *Advances in Neural Information Processing Systems*, vol. 32, 2019, pp. 2396–2405.

[42] C. T. Dinh, N. H. Tran, M. N. H. Nguyen, C. S. Hong, W. Bao, A. Y. Zo5a, and V. Gramoli, "Federated Learning Over Wireless Networks: Convergence Analysis and Resource Allocation," *IEEE/ACM Transactions on Networking*, vol. 29, no. 1, 2021, pp. 398–409.

[43] J. Dong, Y. Shi, and Z. Ding, "Blind over-the-air Computation and Data Fusion via Provable Wirtinger Flow," *IEEE Transactions on Signal Processing*, vol. 68, 2020, pp. 1136–1151.

[44] R. Du, S. Magnusson, and C. Fischione, "The Internet of Things as a Deep Neural Network," *IEEE Communications Magazine*, vol. 58, no. 9, 2020, pp. 20–25.

[45] Z. Du, C. Wu, T. Yoshinaga, K. .-. A. Yau, Y. Ji, and J. Li, "Federated Learning for Vehicular Internet of Things: Recent Advances and Open Issues," *IEEE Open Journal of the Computer Society*, 2020, pp. 45–61.

[46] A. M. Elbir, B. Soner, and S. Coleri, "Federated Learning in Vehicular Networks," *arXiv abs/2006.01412*, 2020, URL: http://arxiv.org/abs/2006.01412.

[47] A. Elgabli, J. Park, C. B. Issaid, and M. Bennis, "Harnessing Wireless Channels for Scalable and Privacy-Preserving Federated Learning," *IEEE Transactions on Communications*, 2021.

[48] K. Elkhalil, A. Hasan, J. Ding, S. Farsiu, and V. Tarokh, "Fisher Auto-Encoders," in *International Conference on Artificial Intelligence and Statistics*, PMLR, pp. 352–360, 2021.

[49] *Ericsson Mobility Report*, 2019, URL: https://www.ericsson.com/en/press-releases/2019/6/ericsson-mobility-report-5g-uptake-even-faster-than-expected.

[50] D. Fan, X. Yuan, and Y.-J. A. Zhang, "Temporal-Structure-Assisted Gradient Aggregation for Over-the-Air Federated Edge Learning," *arXiv abs/2103.02270*, 2021.

[51] C. Finn, P. Abbeel, and S. Levine, "Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks," in *International conference on machine learning*, PMLR, pp. 1126–1135, 2017.

[52] M. Fredrikson, S. Jha, and T. Ristenpart, "Model Inversion Attacks That Exploit Confidence Information and Basic Countermeasures," in *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security*, pp. 1322–1333, 2015.

[53] C. Fung, C. J. Yoon, and I. Beschastnikh, "Mitigating Sybils in Federated Learning Poisoning," *arXiv abs/1808.04866*, 2018.

[54] T. Gafni, N. Shlezinger, K. Cohen, Y. C. Eldar, and H. V. Poor, "Federated Learning: A Signal Processing Perspective," *arXiv abs/2103.17150*, 2021.

[55] J. Goetz, K. Malik, D. Bui, S. Moon, H. Liu, and A. Kumar, "Active Federated Learning," *arXiv abs/1909.12641*, 2019.

[56] M. Goldenbaum and S. Stanczak, "Robust Analog Function Computation via Wireless Multiple-Access Channels," *IEEE Transactions on Communications*, vol. 61, no. 9, 2013, pp. 3863–3877.

[57] M. Goldenbaum and S. Stanczak, "On the Channel Estimation Effort for Analog Computation over Wireless Multiple-Access Channels," *IEEE Wireless Communications Letters*, vol. 3, no. 3, 2014, pp. 261–264.

[58] A. J. Goldsmith and S.-G. Chua, "Adaptive Coded Modulation for Fading Channels," *IEEE Transactions on Communications*, vol. 46, no. 5, 1998, pp. 595–602.

[59] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.

[60] W. Guan, H. Zhang, and V. C. Leung, "Customized Slicing for 6G: Enforcing Artificial Intelligence on Resource Management," *IEEE Network*, 2021, pp. 264–271.

[61] S. Ha, J. Zhang, O. Simeone, and J. Kang, "Coded Federated Computing in Wireless Networks With Straggling Devices and Imperfect CSI," in *Proceedings of the IEEE International Symposium on Information Theory (ISIT)*, IEEE, pp. 2649–2653, 2019.

[62] R. Hamdi, M. Chen, A. B. Said, M. Qaraqe, and H. V. Poor, "Federated Learning Over Energy Harvesting Wireless Networks," *IEEE Internet of Things Journal*, vol. 9, no. 1, 2022, pp. 93–103.

[63] A. Hard, K. Rao, R. Mathews, S. Ramaswamy, F. Beaufays, S. Augenstein, H. Eichner, C. Kiddon, and D. Ramage, "Federated Learning for Mobile Keyboard Prediction," *arXiv abs/1811.03604*, 2018.

[64] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, ser. Springer series in statistics. Springer, 2009.

[65] S. Hayat, E. Yanmaz, and R. Muzaffar, "Survey on Unmanned Aerial Vehicle Networks for Civil Applications: A Communications Viewpoint," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 4, 2016, pp. 2624–2661.

[66] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778, 2016.

[67] H. Hellström, V. Fodor, and C. Fischione, "Over-the-Air Federated Learning with Retransmissions," in *Proceedings of the IEEE 22nd International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, IEEE, pp. 1–5, 2021.

[68] H. Hellström, V. Fodor, and C. Fischione, "Over-the-Air Federated Learning with Retransmissions (Extended Version)," *arXiv abs/2111.10267*, 2021.

[69] G. Hinton, O. Vinyals, and J. Dean, "Distilling the Knowledge in a Neural Network," *arXiv abs/1503.02531*, 2015.

[70] Y. Hu, M. Chen, M. Chen, Z. Yang, M. Shikh-Bahaei, H. V. Poor, and S. Cui, "Energy Minimization for Federated Learning with IRS-Assisted Over-the-Air Computation," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Toronto, ON, Canada, 2021.

[71] K. Huang, G. Zhu, C. You, J. Zhang, Y. Du, and D. Liu, "Communication, Computing, and Learning on the Edge," in *Proceedings of the 2018 IEEE International Conference on Communication Systems (ICCS)*, IEEE, pp. 268–273, 2018.

[72] S.-J. Huang and Z.-H. Zhou, "Active Query Driven by Uncertainty and Diversity for Incremental Multi-Label Learning," in *2013 IEEE 13th International Conference on Data Mining*, IEEE, pp. 1079–1084, 2014.

[73] F. Hussain, S. A. Hassan, R. Hussain, and E. Hossain, "Machine Learning for Resource Management in Cellular and IoT Networks: Potentials, Current Solutions, and Open Challenges," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 2, 2020, pp. 1251–1275.

[74] A. Imteaj, U. Thakker, S. Wang, J. Li, and M. H. Amini, "A Survey on Federated Learning for Resource-Constrained IoT Devices," *IEEE Internet of Things Journal*, 2021, pp. 1–24.

[75] Y. Inagaki, R. Shinkuma, T. Sato, and E. Oki, "Prioritization of Mobile IoT Data Transmission Based on Data Importance Extracted From Machine Learning Model," *IEEE Access*, vol. 7, 2019, pp. 93 611–93 620.

[76] S. R. Islam, M. Zeng, O. A. Dobre, and K.-S. Kwak, "Nonorthogonal Multiple Access (NOMA): How It Meets 5G and Beyond," *Wiley 5G Ref: The Essential 5G Reference Online*, 2019, pp. 1–28.

[77] M. Jaggi, V. Smith, M. Takác, J. Terhorst, S. Krishnan, T. Hofmann, and M. I. Jordan, "Communication-efficient Distributed Dual Coordinate Ascent," in *Advances in Neural Information Processing Systems*, pp. 3068–3076, 2014.

[78] Y.-S. Jeon, M. M. Amiri, and N. Lee, "Communication-Efficient Federated Learning over MIMO Multiple Access Channels," 2022, under review.

[79] Y.-S. Jeon, M. M. Amiri, J. Li, and H. V. Poor, "A Compressive Sensing Approach for Federated Learning over Massive MIMO Communication Systems," *IEEE Transactions on Wireless Communications*, vol. 20, no. 3, 2021, pp. 1990–2004.

[80] E. Jeong, S. Oh, H. Kim, J. Park, M. Bennis, and S.-L. Kim, "Communication-Efficient On-Device Machine Learning: Federated Distillation and Augmentation Under Non-i.i.d. Private Data," *arXiv abs/1811.11479*, 2018.

[81] R. Jiang and S. Zhou, "Cluster-Based Cooperative Digital Over-the-Air Aggregation for Wireless Federated Edge Learning," in *Proceedings of the 2020 IEEE/CIC International Conference on Communications in China (ICCC)*, IEEE, pp. 887–892, 2020.

[82] T. Jiang and Y. Shi, "Over-the-Air Computation via Intelligent Reflecting Surfaces," in *Proceedings of the 2019 IEEE Global Communications Conference (GLOBECOM)*, IEEE, pp. 1–6, 2019.

[83] R. Jin, X. He, and H. Dai, "Communication Efficient Federated Learning with Energy Awareness over Wireless Networks," *IEEE Transactions on Wireless Communications*, 2022, Early Access.

[84] P. Kairouz, H. B. McMahan, B. Avent, A. Bellet, M. Bennis, A. N. Bhagoji, K. Bonawitz, Z. Charles, G. Cormode, R. Cummings, R. G. L. D'Oliveira, S. E. Rouayheb, D. Evans, J. Gardner, Z. A. Garrett, A. Gascón, B. Ghazi, P. B. Gibbons, M. Gruteser, Z. Harchaoui, C. He, L. He, Z. Huo, B. Hutchinson, J. Hsu, M. Jaggi, T. Javidi, G. Joshi, M. Khodak, J. Konečný, A. Korolova, F. Koushanfar, O. Koyejo, T. Lepoint, Y. Liu, P. Mittal, M. Mohri, R. Nock, A. Özgür, R. Pagh, M. Raykova, H. Qi, D. Ramage, R. Raskar, D. X. Song, W. Song, S. U. Stich, Z. Sun, A. T. Suresh, F. Tramèr, P. Vepakomma, J. Wang, L. Xiong, Z. Xu, Q. Yang, F. X. Yu, H. Yu, and S. Zhao, "Advances and Open Problems in Federated Learning," *Foundations and Trends® in Machine Learning*, vol. 14, no. 1-2, 2021, pp. 1–210.

[85] Y. Koda, K. Yamamoto, T. Nishio, and M. Morikura, "Differentially Private Aircomp Federated Learning with Power Adaptation Harnessing Receiver Noise," in *Proceedings of the IEEE Global Communications Conference*, pp. 1–6, Taipei, Taiwan, 2021.

[86] J. Konečný, H. B. McMahan, D. Ramage, and P. Richtárik, "Federated Optimization: Distributed Machine Learning for On-Device Intelligence," *arXiv abs/1610.02527*, 2016.

[87] A. Krizhevsky, *Learning Multiple Layers of Features from Tiny Images*, 2009, URL: https://www.cs.toronto.edu/~kriz/learning-features-2009-TR.pdf.

[88] B. Lakshminarayanan, A. Pritzel, and C. Blundell, "Simple and Scalable Predictive Uncertainty Estimation Using Deep Ensembles," *Advances in Neural Information Processing Systems*, vol. 30, 2017.

[89] Y. LeCun, L. Bottou, Y. Bengio, P. Haffner, *et al.*, "Gradient-Based Learning Applied to Document Recognition," *Proceedings of the IEEE*, vol. 86, no. 11, 1998, pp. 2278–2324.

[90] C.-S. Lee, M.-H. Wang, S.-J. Yen, T.-H. Wei, I.-C. Wu, P.-C. Chou, C.-H. Chou, M.-W. Wang, and T.-H. Yan, "Human vs. Computer Go: Review and Prospect [Discussion Forum]," *IEEE Computational intelligence magazine*, vol. 11, no. 3, 2016, pp. 67–72.

[91] J. Leng, Z. Lin, M. Ding, P. Wang, D. Smith, and B. Vucetic, "Client Scheduling in Wireless Federated Learning Based on Channel and Learning Qualities," *IEEE Wireless Communications Letters*, 2022.

[92] T. Li, A. K. Sahu, A. Talwalkar, and V. Smith, "Federated Learning: Challenges, Methods, and Future Directions," *IEEE Signal Processing Magazine*, vol. 37, no. 3, 2020, pp. 50–60.

[93] T. Li, A. K. Sahu, M. Zaheer, M. Sanjabi, A. Talwalkar, and V. Smith, "Federated Optimization in Heterogeneous Networks," in *Proceedings of Machine Learning and Systems*, 2020, pp. 429–450.

[94] T. Li, M. Sanjabi, A. Beirami, and V. Smith, "Fair Resource Allocation in Federated Learning," *arXiv abs/1905.10497*, 2019.

[95] Y. Lin, S. Han, H. Mao, Y. Wang, and W. J. Dally, "Deep Gradient Compression: Reducing the Communication Bandwidth for Distributed Training," *arXiv abs/1712.01887*, 2017.

[96] D. Liu and O. Simeone, "Privacy for Free: Wireless Federated Learning via Uncoded Transmission With Adaptive Power Control," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 1, 2021, pp. 170–185.

[97] D. Liu and O. Simeone, "Channel-driven Monte Carlo Sampling for Bayesian Distributed Learning in Wireless Data Centers," *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 2, 2022, pp. 562–577.

[98] D. Liu, G. Zhu, Q. Zeng, J. Zhang, and K. Huang, "Wireless Data Acquisition for Edge Learning: Data-Importance Aware Retransmission," *IEEE Transactions on Wireless Communications*, vol. 20, no. 1, 2020, pp. 406–420.

[99] D. Liu, G. Zhu, J. Zhang, and K. Huang, "Wireless Data Acquisition for Edge Learning: Importance-Aware Retransmission," in *Proceedings of the 2019 IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, IEEE, pp. 1–5, 2019.

[100] D. Liu, G. Zhu, J. Zhang, and K. Huang, "Data-Importance Aware User Scheduling for Communication-efficient Edge Machine Learning," *IEEE Transactions on Cognitive Communications and Networking*, vol. 7, no. 1, 2020, pp. 265–278.

[101] H. Liu, X. Yuan, and Y.-J. A. Zhang, "CSIT-Free Federated Edge Learning via Reconfigurable Intelligent Surface," *arXiv abs/1905.10497*, 2021.

[102] H. Liu, X. Yuan, and Y.-J. A. Zhang, "Reconfigurable Intelligent Surface Enabled Federated Learning: A Unified Communication-Learning Design Approach," *IEEE Transactions on Wireless Communications*, 2021, pp. 7595–7609.

[103] W. Liu, X. Zang, Y. Li, and B. Vucetic, "Over-the-Air Computation Systems: Optimization, Analysis and Scaling Laws," *IEEE Transactions on Wireless Communications*, vol. 19, no. 8, 2020, pp. 5488–5502.

[104] M. M. Amiri, D. Gündüz, S. R. Kulkarni, and H. V. Poor, "Federated Learning With Quantized Global Model Updates," *arXiv abs/2006.10672*, 2020.

[105] M. M. Amiri, S. R. Kulkarni, and H. V. Poor, "Federated Learning With Downlink Device Selection," in *Proceedings of the IEEE International Workshop on Signal Processing Advances in Wireless Communications*, Lucca, Italy, 2021.

[106] M. M. Amiri, T. M. Duman, and D. Gündüz, "Collaborative Machine Learning at the Wireless Edge with Blind Transmitters," in *Proceedings of the IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, pp. 1–5, Ottawa, ON, Canada, 2019.

[107] C. Ma, J. Konečný, M. Jaggi, V. Smith, M. I. Jordan, P. Richtárik, and M. Takáč, "Distributed Optimization with Arbitrary Local Solvers," *Optimization Methods and Software*, vol. 32, no. 4, 2017, pp. 813–848.

[108] M. A. Maddah-Ali and U. Niesen, "Fundamental Limits of Caching," *IEEE Transactions on Information Theory*, vol. 60, no. 5, 2014, pp. 2856–2867.

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

[110] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient Learning of Deep Networks from Decentralized Data," in *Artificial intelligence and Statistics*, PMLR, pp. 1273–1282, 2017.

[111] L. Melis, C. Song, E. D. Cristofaro, and V. Shmatikov, "Exploiting Unintended Feature Leakage in Collaborative Learning," in *Proceedings of the IEEE Symposium on Security and Privacy*, pp. 691–706, San Francisco, CA, USA, 2019.

[112] B. Nazer and M. Gastpar, "Computation over Multiple-Access Channels," *IEEE Transactions on Information Theory*, vol. 53, no. 10, 2007, pp. 3498–3516.

[113] A. Nedić, A. Olshevsky, and M. G. Rabbat, "Network Topology and Communication-Computation Tradeoffs in Decentralized Optimization," *Proceedings of the IEEE*, vol. 106, no. 5, 2018, pp. 953–976.

[114] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free Massive MIMO Versus Small Cells," *IEEE Transactions on Wireless Communications*, vol. 16, no. 3, 2017, pp. 1834–1850.

[115] D. C. Nguyen, M. Ding, P. N. Pathirana, A. Seneviratne, J. Li, D. Niyato, and H. V. Poor, "Federated Learning for Industrial Internet of Things in Future Industries," *arXiv abs/2105.14659*, 2021.

[116] S. Niknam, H. S. Dhillon, and J. H. Reed, "Federated Learning for Wireless Communications: Motivation, Opportunities, and Challenges," *IEEE Communications Magazine*, vol. 58, no. 6, 2020, pp. 46–51.

[117] H. Ning, H. Wang, Y. Lin, W. Wang, S. Dhelim, F. Farha, J. Ding, and M. Daneshmand, "A Survey on Metaverse: the State-of-the-art, Technologies, Applications, and Challenges," *arXiv abs/2111.09673*, 2021.

[118] T. Nishio and R. Yonetani, "Client Selection for Federated Learning with Heterogeneous Resources in Mobile Edge," in *Proceedings of the 2019 IEEE International Conference on Communications (ICC)*, IEEE, pp. 1–7, 2019.

[119] S. Oh, J. Park, E. Jeong, H. Kim, M. Bennis, and S.-L. Kim, "Mix2FLD: Downlink Federated Learning After Uplink Federated Distillation With Two-Way Mixup," *IEEE Communications Letters*, vol. 24, no. 10, 2020, pp. 2211–2215.

[120] J. Park, S. Samarakoon, M. Bennis, and M. Debbah, "Wireless Network Intelligence at the Edge," *Proceedings of the IEEE*, vol. 107, no. 11, 2019, pp. 2204–2239.

[121] J. Park, S. Wang, A. Elgabli, S. Oh, E. Jeong, H. Cha, H. Kim, S.-L. Kim, and M. Bennis, "Distilling on-Device Intelligence at the Network Edge," *arXiv abs/1908.05895*, 2019.

[122] S. Park, S. Jung, H. Lee, J. Kim, and J.-H. Kim, "Large-Scale Water Quality Prediction Using Federated Sensing and Learning: A Case Study with Real-World Sensing Big-Data," *Sensors*, vol. 21, no. 4, 2021, p. 1462.

[123] M. Półka, S. Ptak, and Ł. Kuziora, "The Use of UAV's for Search and Rescue Operations," *Procedia Engineering*, vol. 192, 2017, pp. 748–752.

[124] S. Prakash, H. Hashemi, Y. Wang, M. Annavaram, and S. Avestimehr, "Byzantine-Resilient Federated Learning with Heterogeneous Data Distribution," *arXiv abs/2010.07541*, 2020.

[125] S. Pu and A. Nedić, "Distributed Stochastic Gradient Tracking Methods," *Mathematical Programming*, vol. 187, no. 1, 2020, pp. 409–457.

[126] Z. Qin, G. Y. Li, and H. Ye, "Federated Learning and Wireless Communications," *IEEE Wireless Communications*, vol. 28, no. 5, 2021.

[127] A. Reisizadeh, A. Mokhtari, H. Hassani, A. Jadbabaie, and R. Pedarsani, "FedPAQ: A Communication-Efficient Federated Learning Method with Periodic Averaging and Quantization," in *Proceedings of the International Conference on Artificial Intelligence and Statistics*, PMLR, pp. 2021–2031, 2020.

[128] J. Ren, Y. He, D. Wen, G. Yu, K. Huang, and D. Guo, “Scheduling for Cellular Federated Edge Learning With Importance and Channel Awareness,” *IEEE Transactions on Wireless Communications*, vol. 19, no. 11, 2020, pp. 7690–7703.

[129] J. Ren, G. Yu, and G. Ding, “Accelerating DNN Training in Wireless Federated Edge Learning Systems,” *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 1, 2020, pp. 219–232.

[130] P. Rost, C. Mannweiler, D. S. Michalopoulos, C. Sartori, V. Sciancalepore, N. Sastry, O. Holland, S. Tayade, B. Han, D. Bega, *et al.*, “Network Slicing to Enable Scalability and Flexibility in 5G Mobile Networks,” *IEEE Communications Magazine*, vol. 55, no. 5, 2017, pp. 72–79.

[131] W. Saad, M. Bennis, and M. Chen, “A Vision of 6G Wireless Systems: Applications, Trends, Technologies, and Open Research Problems,” *IEEE Network*, vol. 34, no. 3, 2019, pp. 134–142.

[132] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A. Li, and K. Higuchi, “Non-Orthogonal Multiple Access (NOMA) for Cellular Future Radio Access,” in *Proceedings of the 2013 IEEE 77th Vehicular Technology Conference (VTC Spring)*, IEEE, pp. 1–5, 2013.

[133] M. Salehi and E. Hossain, “Federated Learning in Unreliable and Resource-Constrained Cellular Wireless Networks,” *IEEE Transactions on Communications*, 2021.

[134] S. Samarakoon, M. Bennis, W. Saad, and M. Debbah, “Distributed Federated Learning for Ultra-Reliable Low-Latency Vehicular Communications,” *IEEE Transactions on Communications*, vol. 68, no. 2, 2019, pp. 1146–1159.

[135] F. Seide, H. Fu, J. Droppo, G. Li, and D. Yu, “1-bit Stochastic Gradient Descent and its Application to Data-Parallel Distributed Training of Speech DNNs,” in *Proceedings of the 15th Annual Conference of the International Speech Communication Association*, 2014.

[136] M. Seif, R. Tandon, and M. Li, “Wireless Federated Learning with Local Differential Privacy,” in *Proceedings of the IEEE International Symposium on Information Theory*, pp. 2604–2609, 2020.

[137] T. Sery and K. Cohen, "A Sequential Gradient-Based Multiple Access for Distributed Learning over Fading Channels," in *Proceedings of the 2019 57th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, IEEE, pp. 303–307, 2019.

[138] T. Sery and K. Cohen, "On Analog Gradient Descent Learning over Multiple Access Fading Channels," *IEEE Transactions on Signal Processing*, vol. 68, 2020, pp. 2897–2911.

[139] B. Settles, "Active Learning Literature Survey," University of Wisconsin-Madison Department of Computer Sciences, Tech. Rep., 2009.

[140] B. Settles, "Active Learning," *Synthesis Lectures on Artificial Intelligence and Machine Learning*, vol. 6, no. 1, 2012, pp. 1–114.

[141] O. Shamir, N. Srebro, and T. Zhang, "Communication-Efficient Distributed Optimization Using an Approximate Newton-Type Method," in *Proceedings of the International Conference on Machine Learning*, PMLR, pp. 1000–1008, 2014.

[142] W. Shi, S. Zhou, and Z. Niu, "Device Scheduling with Fast Convergence for Wireless Federated Learning," in *Proceedings of the 2020 IEEE International Conference on Communications (ICC)*, IEEE, pp. 1–6, 2020.

[143] Y. Shi, Y. Zhou, and Y. Shi, "Over-the-Air Decentralized Federated Learning," in *Proceedings of the 2021 IEEE International Symposium on Information Theory (ISIT)*, IEEE, pp. 455–460, 2021.

[144] Y. Shi, K. Yang, T. Jiang, J. Zhang, and K. B. Letaief, "Communication-efficient Edge AI: Algorithms and Systems," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 4, 2020, pp. 2167–2191.

[145] R. Shinkuma and T. Nishio, "Data Assessment and Prioritization in Mobile Networks for Real-Time Prediction of Spatial Information with Machine Learning," in *Proceedings of the IEEE First International Workshop on Network Meets Intelligent Computations (NMIC)*, IEEE, pp. 1–6, 2019.

[146] N. Shlezinger, M. Chen, Y. C. Eldar, H. V. Poor, and S. Cui, "UVeQFed: Universal Vector Quantization for Federated Learning," *IEEE Transactions on Signal Processing*, vol. 69, 2020, pp. 500–514.

[147] H. Sifaou and G. Y. Li, "Robust Federated Learning via Over-The-Air Computation," *arXiv abs/2111.01221*, 2021.

[148] J. M. B. da Silva Jr., K. Ntougias, I. Krikidis, G. Fodor, and C. Fischione, "Simultaneous Wireless Information and Power Transfer for Federated Learning," *arXiv abs/2104.12749*, 2021.

[149] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," *arXiv abs/1409.1556*, 2014.

[150] J. So, B. Güler, and A. S. Avestimehr, "CodedPrivateML: A Fast and Privacy-Preserving Framework for Distributed Machine Learning," *IEEE Journal on Selected Areas in Information Theory*, vol. 2, no. 1, 2021, pp. 441–451.

[151] E. C. Strinati, S. Barbarossa, J. L. Gonzalez-Jimenez, D. Ktenas, N. Cassiau, L. Maret, and C. Dehos, "6G: The Next Frontier: From Holographic Messaging to Artificial Intelligence Using Subterahertz and Visible Light Communication," *IEEE Vehicular Technology Magazine*, vol. 14, no. 3, 2019, pp. 42–50.

[152] C. Sun, A. Shrivastava, S. Singh, and A. Gupta, "Revisiting Unreasonable Effectiveness of Data in Deep Learning Era," in *Proceedings of the IEEE international conference on computer vision*, pp. 843–852, 2017.

[153] Y. Sun, S. Zhou, and D. Gündüz, "Energy-Aware Analog Aggregation for Federated Learning with Redundant Data," in *Proceedings of the 2020 IEEE International Conference on Communications (ICC)*, IEEE, pp. 1–7, 2020.

[154] S. Sundhar Ram, A. Nedić, and V. V. Veeravalli, "Distributed Stochastic Subgradient Projection Algorithms for Convex Optimization," *Journal of optimization theory and applications*, vol. 147, no. 3, 2010, pp. 516–545.

[155] H. Tang, X. Lian, C. Yu, T. Zhang, and J. Liu, "DoubleSqueeze: Parallel Stochastic Gradient Descent with Double-Pass Error-Compensated Compression," in *Proceedings of the International Conference on Machine Learning*, Long Beach, CA, 2019.

[156] *Timing Advance (TA) in LTE*, 2010, URL: http://4g5gworld.com/blog/timing-advance-ta-lte.

[157] S. Timotheou, I. Krikidis, G. Zheng, and B. Ottersten, "Beamforming for MISO Interference Channels with QoS and RF Energy Transfer," *IEEE Transactions on Wireless Communications*, vol. 13, no. 5, 2014, pp. 2646–2658.

[158] N. H. Tran, W. Bao, A. Zo5a, M. N. Nguyen, and C. S. Hong, "Federated Learning over Wireless Networks: Optimization Model Design and Analysis," in *Proceedings of the IEEE INFOCOM 2019 Conference on Computer Communications*, IEEE, pp. 1387–1395, 2019.

[159] A. Vempaty, L. Tong, and P. K. Varshney, "Distributed Inference with Byzantine Data: State-of-the-Art Review on Data Falsification Attacks," *IEEE Signal Processing Magazine*, vol. 30, no. 5, 2013, pp. 65–75.

[160] T. T. Vu, D. T. Ngo, N. H. Tran, H. Q. Ngo, M. N. Dao, and R. H. Middleton, "Cell-Free Massive MIMO for Wireless Federated Learning," *IEEE Transactions on Wireless Communications*, vol. 19, no. 10, 2020, pp. 6377–6392.

[161] M. M. Wadu, S. Samarakoon, and M. Bennis, "Federated Learning Under Channel Uncertainty: Joint Client Scheduling and Resource Allocation," in *Proceedings of the 2020 IEEE Wireless Communications and Networking Conference (WCNC)*, IEEE, pp. 1–6, 2020.

[162] J. Wang, C. Jiang, H. Zhang, Y. Ren, K.-C. Chen, and L. Hanzo, "Thirty Years of Machine Learning: The Road to Pareto-Optimal Wireless Networks," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 3, 2020, pp. 1472–1514.

[163] S. Wang, Y.-C. Wu, M. Xia, R. Wang, and H. V. Poor, "Machine Intelligence at the Edge With Learning Centric Power Allocation," *IEEE Transactions on Wireless Communications*, vol. 19, no. 11, 2020, pp. 7293–7308.

[164] X. Wang, Y. Han, V. C. Leung, D. Niyato, X. Yan, and X. Chen, "Convergence of Edge Computing and Deep Learning: A Comprehensive Survey," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 2, 2020, pp. 869–904.

[165] Z. Wang, J. Qiu, Y. Zhou, Y. Shi, L. Fu, W. Chen, and K. B. Letaief, "Federated Learning via Intelligent Reflecting Surface," *IEEE Transactions on Wireless Communications*, 2021, pp. 808–822.

[166] K. Wei, J. Li, M. Ding, C. Ma, H. H. Yang, F. Farokhi, S. Jin, T. Q. Quek, and H. V. Poor, "Federated Learning with Differential Privacy: Algorithms and Performance Analysis," *IEEE Transactions on Information Forensics and Security*, vol. 15, 2020, pp. 3454–3469.

[167] D. Wen, X. Li, Q. Zeng, J. Ren, and K. Huang, “An Overview of Data-Importance Aware Radio Resource Management for Edge Machine Learning,” *Journal of Communications and Information Networks*, vol. 4, no. 4, 2019, pp. 1–14.

[168] A. G. Wilson and P. Izmailov, “Bayesian Deep Learning and a Probabilistic Perspective of Generalization,” *Advances in neural information processing systems*, vol. 33, 2020, pp. 4697–4708.

[169] T. Wu, F. Wu, J.-M. Redoute, and M. R. Yuce, “An Autonomous Wireless Body Area Network Implementation Towards IoT Connected Healthcare Applications,” *IEEE Access*, vol. 5, 2017, pp. 11 413–11 422.

[170] Q. Xia, W. Ye, Z. Tao, J. Wu, and Q. Li, “A Survey of Federated Learning for Edge Computing: Research Problems and Solutions,” *High-Confidence Computing*, 2021, p. 100 008.

[171] W. Xia, W. Wen, K. .-. Wong, T. Q. S. Quek, J. Zhang, and H. Zhu, “Federated-Learning-Based Client Scheduling for Low-Latency Wireless Communications,” *IEEE Wireless Communications*, vol. 28, no. 2, 2021, pp. 32–38.

[172] R. Xin, S. Kar, and U. A. Khan, “Decentralized Stochastic Optimization and Machine Learning: A Unified Variance-Reduction Framework for Robust Performance and Fast Convergence,” *IEEE Signal Processing Magazine*, vol. 37, no. 3, 2020, pp. 102–113.

[173] H. Xing, O. Simeone, and S. Bi, “Decentralized Federated Learning via SGD over Wireless D2D Networks,” in *Proceedings of the 2020 IEEE 21st International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, IEEE, pp. 1–5, 2020.

[174] J. Xu and H. Wang, “Client Selection and Bandwidth Allocation in Wireless Federated Learning Networks: A Long-Term Perspective,” *IEEE Transactions on Wireless Communications*, vol. 20, no. 2, 2021, pp. 1188–1200.

[175] J. Xu, H. Wang, and L. Chen, “Bandwidth Allocation for Multiple Federated Learning Services in Wireless Edge Networks,” *arXiv abs/2101.03627*, 2021.

[176] P. Xue, P. Gong, J. H. Park, D. Park, and D. K. Kim, "Max-Min Fairness Based Radio Resource Management in Fourth Generation Heterogeneous Networks," in *Proceedings of the 9th International Symposium on Communications and Information Technology*, IEEE, pp. 208–213, 2009.

[177] P. Yagol, F. Ramos, S. Trilles, J. Torres-Sospedra, and F. J. Perales, "New Trends in Using Augmented Reality Apps for Smart City Contexts," *ISPRS International Journal of Geo-Information*, vol. 7, no. 12, 2018, p. 478.

[178] H. H. Yang, A. Arafa, T. Q. S. Quek, and H. V. Poor, "Age-Based Scheduling Policy for Federated Learning in Mobile Edge Networks," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2020.

[179] K. Yang, T. Jiang, Y. Shi, and Z. Ding, "Federated Learning via Over-the-Air Computation," *IEEE Transactions on Wireless Communications*, vol. 19, no. 3, 2020, pp. 2022–2035.

[180] Q. Yang, Y. Liu, Y. Cheng, Y. Kang, T. Chen, and H. Yu, *Federated Learning*, ser. Synthesis Lectures on Artificial Intelligence and Machine Learning. Morgan & Claypool Publishers, 2019.

[181] Z. Yang, M. Chen, W. Saad, C. S. Hong, and M. Shikh-Bahaei, "Energy Efficient Federated Learning Over Wireless Communication Networks," *IEEE Transactions on Wireless Communications*, vol. 20, no. 3, 2021, pp. 1935–1949.

[182] W. Yu and J. M. Cioffi, "On Constant Power Water-Filling," in *IEEE International Conference on Communications. Conference Record (ICC)*, IEEE, vol. 6, pp. 1665–1669, 2002.

[183] S. Yue, J. Ren, J. Xin, D. Zhang, Y. Zhang, and W. Zhuang, "Efficient Federated Meta-Learning over Multi-Access Wireless Networks," *IEEE Journal on Selected Areas in Communications*, vol. 40, 2022, pp. 1556–1570.

[184] X. Zang, W. Liu, Y. Li, and B. Vucetic, "Over-the-Air Computation Systems: Optimal Design with Sum-Power Constraint," *IEEE Wireless Communications Letters*, vol. 9, no. 9, 2020, pp. 1524–1528.

[185] Q. Zeng, Y. Du, and K. Huang, "Wirelessly Powered Federated Edge Learning: Optimal Tradeoffs Between Convergence and Power Transfer," *arXiv abs/2102.12357*, 2021.

[186] Q. Zeng, Y. Du, K. Huang, and K. K. Leung, "Energy-Efficient Radio Resource Allocation for Federated Edge Learning," in *Proceedings of the 2020 IEEE International Conference on Communications Workshops (ICC Workshops)*, IEEE, pp. 1–6, 2020.

[187] T. Zeng, O. Semiari, M. Chen, W. Saad, and M. Bennis, "Federated Learning on the Road: Autonomous Controller Design for Connected and Autonomous Vehicles," *arXiv abs/2102.03401*, 2021.

[188] H. Zhang, M. Cisse, Y. N. Dauphin, and D. Lopez-Paz, "mixup: Beyond Empirical Risk Minimization," *arXiv abs/1710.09412*, 2017.

[189] N. Zhang and M. Tao, "Gradient Statistics Aware Power Control for Over-the-Air Federated Learning," *IEEE Transactions on Wireless Communications*, 2021, pp. 5115–5128.

[190] S. Zhang, S. C. Liew, and P. P. Lam, "Hot Topic: Physical-Layer Network Coding," in *Proceedings of the 12th Annual International Conference on Mobile Computing and Networking*, ACM, pp. 358–365, 2006.

[191] Z. Zhang, Y. Xiao, Z. Ma, M. Xiao, Z. Ding, X. Lei, G. K. Karagiannidis, and P. Fan, "6G Wireless Networks: Vision, Requirements, Architecture, and Key Technologies," *IEEE Vehicular Technology Magazine*, vol. 14, no. 3, 2019, pp. 28–41.

[192] J. Zhao, "A Survey of Intelligent Reflecting Surfaces (IRSs): Towards 6G Wireless Communication Networks," *arXiv abs/1907.04789*, 2019.

[193] Y. Zhao, M. Li, L. Lai, N. Suda, D. Civin, and V. Chandra, "Federated Learning With non-IID Data," *arXiv abs/1806.00582*, 2018.

[194] Z. Zhou, X. Chen, E. Li, L. Zeng, K. Luo, and J. Zhang, "Edge Intelligence: Paving the Last Mile of Artificial Intelligence with Edge Computing," *Proceedings of the IEEE*, vol. 107, no. 8, 2019, pp. 1738–1762.

[195] G. Zhu, L. Chen, and K. Huang, "Over-the-Air Computation in MIMO Multi-Access Channels: Beamforming and Channel Feedback," *CoRR, vol. abs/1803.11129*, 2018.

[196] G. Zhu, Y. Du, D. Gündüz, and K. Huang, "One-Bit Over-the-Air Aggregation for Communication-Efficient Federated Edge Learning: Design and Convergence Analysis," *IEEE Transactions on Wireless Communications*, vol. 20, no. 3, 2020, pp. 2120–2135.

[197] G. Zhu, D. Liu, Y. Du, C. You, J. Zhang, and K. Huang, "Towards an Intelligent Edge: Wireless Communication Meets Machine Learning," *arXiv abs/1809.00343*, 2018.

[198] G. Zhu, D. Liu, Y. Du, C. You, J. Zhang, and K. Huang, "Toward an Intelligent Edge: Wireless Communication Meets Machine Learning," *IEEE Communications Magazine*, vol. 58, no. 1, 2020, pp. 19–25.

[199] G. Zhu, Y. Wang, and K. Huang, "Broadband Analog Aggregation for Low-Latency Federated Edge Learning (extended version)," *arXiv abs/1812.11494*, 2018.

[200] G. Zhu, Y. Wang, and K. Huang, "Broadband Analog Aggregation for Low-Latency Federated Edge Learning," *IEEE Transactions on Wireless Communications*, vol. 19, no. 1, 2020, pp. 491–506.

[201] G. Zhu, J. Xu, K. Huang, and S. Cui, "Over-the-Air Computing for Wireless Data Aggregation in Massive IoT," *IEEE Wireless Communications*, vol. 28, no. 4, 2021, pp. 57–65.

[202] M. Zhu and S. Gupta, "To Prune, or not to Prune: Exploring the Efficacy of Pruning for Model Compression," *arXiv abs/1710.01878*, 2017.

[203] M. Zinkevich, M. Weimer, L. Li, and A. J. Smola, "Parallelized Stochastic Gradient Descent," in *Advances in Neural Information Processing Systems*, pp. 2595–2603, 2010.