The Scanning the Literature column offers succinct overviews of recently published papers in the networking field. Each summary outlines the paper's primary concept, methodology, and technical contributions. The column's objective is to present the latest advancements in networking research to the readers of *IEEE Network* magazine. Authors are also encouraged to recommend their newly published work for inclusion in the column, with particular appreciation for papers featuring innovative ideas, robust research, and significant contributions to the field. Authors interested in having their papers featured in the column should contact Xiaohua Tian at the email address provided below.

Xiaohua Tian,
School of Electronic Information and Electrical Engineering,
Shanghai Jiao Tong University
xtian@sjtu.edu.cn

In the Internet of Things (IoT) era, the number of IoT nodes is expected to reach billions, with deployment densities exceeding a million per square kilometer. This surge in connectivity results in the generation of vast datasets. However, not all raw data collected in certain applications possess inherent value. In some cases, our interest may lie in specific metrics such as averages or extremities. Given that sensor nodes are typically resource-constrained, they often need to transmit raw data to central nodes for computation, following a communication-then-computation paradigm. In situations with extensive device deployments, this method can introduce significant latency. Therefore, the concept of over-the-air computation has been introduced to seamlessly integrate communication and computation, effectively performing computations during the data transmission process. This column aims to provide an academic overview of the latest advancements in over-the-air computation. We will discuss recent developments in over-the-air computation, including its combination with reconfigurable intelligent surfaces (RIS), power control, blind estimation, energy optimization, beamforming design, and its application in convolutional neural networks (CNN).

Over-the-air computation is a groundbreaking technology that leverages the waveform superposition characteristics of multi-path channels to enable swift wireless data aggregation in IoT networks. However, the performance of over-the-air computation faces constraints due to suboptimal channel conditions prevalent across all links between IoT devices and access points. As a result, some research efforts explore the integration of reconfigurable intelligent surfaces (RIS) into the context of over-the-air computation. RIS deployment involves the adaptive manipulation of propagation channels, effectively enhancing received signal power and alleviating performance bottlenecks. The following paper demonstrates how to utilize the state-of-the-art in RIS to assist over-the-air computation.

Over-the-Air Computation via Reconfigurable Intelligent Surface

W. Fang, Y. Jiang, Y. Shi, Y. Zhou, W. Chen, and K. B. Letaief, "Over-the-air computation via reconfigurable intelligent surface," *IEEE Trans. Commun.*, vol. 69, no. 12, pp. 8612–8626, Dec. 2021.

In this paper, Fang et al. propose a reconfigurable RIS-assisted over-the-air computation (AirComp) system aimed at enhancing wireless data fusion in IoT networks. The authors devise a joint design of AirComp transceivers and RIS phase shifts to minimize distortion. They propose an alternating minimization method within the successive convex approximation (SCA) framework to solve the non-convex problem. The problem is decomposed into two subproblems, and convex approximations are constructed using SCA. The proposed algorithm, named AlterMinSCA, alternately solves these subproblems until convergence. To decrease computational complexity, they transform the subproblem into a smooth convex-concave saddle point problem and propose a Mirror-Prox method for closed-form updates. The proposed algorithm achieves comparable distortion performance to state-of-the-art algorithms but with significantly reduced computation time. Simulation results validate the effectiveness of the proposed algorithm in enhancing the performance of RIS-assisted AirComp systems. The algorithm outperforms other baseline methods in terms of both mean square error (MSE) and computation time. The authors conclude that their proposed algorithm holds potential for high-density RIS-assisted AirComp systems.

Digital Object Identifier:
10.1109/MNET.2023.3332537

AirComp utilizes the waveform superposition property of signals to facilitate rapid data fusion. However, signal distortion in fading channel transmission can lead to waveform deformities, resulting in computational inaccuracies. Typically, AirComp uses a single time slot for data fusion, ensuring high efficiency. However, to maintain unbiased data fusion, each IoT node must pre-amplify its data to ensure uniform signal amplitudes at the receiver. Pre-amplification often employs the channel inversion principle (power inversely proportional to channel coefficients), which poses high power demands in deep fading transmission. However, accurately acquiring channel coefficients from multiple nodes remains a challenging task. As a result, some research efforts consider using multiple time slots for data transmission, leveraging the time-varying channel gains to achieve power control. In the following paper, Tang et al. propose a multi-slot over-the-air computation (MS-AirComp) framework for sum estimation in fading channels.

Multi-Slot Over-the-Air Computation in Fading Channels

S. Tang, P. Popovski, C. Zhang, and S. Obana, "Multi-slot over-the-air computation in fading channels," *IEEE Trans. Wireless Commun.*, vol. 22, no. 10, pp. 6766–6777, Oct. 2023.

This paper presents a multi-slot over-the-air computation (MS-AirComp) framework designed to address the issue of signal distortion in fading channels in IoT systems. The framework aims to enhance channel gains and facilitate power control by distributing signal transmission across multiple slots and selecting slots with high channel gains for transmission. The authors provide a theoretical analysis of the computation error and derive optimal parameters for the framework. To further improve the performance of the MS-AirComp framework, the paper introduces a method called SelFirst. This method selects the slot with the highest channel gain for transmission, reducing computation mean square error (MSE) and transmission power. Simulations are conducted to evaluate the effectiveness of the proposed methods and compare them with existing approaches. The results demonstrate that the MS-AirComp framework and SelFirst method achieve an appropriate tradeoff between computation MSE and transmission power. The authors highlight the scalability of the proposed method with the number of nodes and emphasize the potential of the MS-AirComp framework for mitigating signal distortion in wireless networks. The paper also discusses future work, including addressing signal misalignment and enhancing system performance with multiple antennas.

As previously mentioned, exact channel state information (CSI) is required for AirComp. However, accurate CSI collection itself is challenging for each node. Currently, the primary constraint of AirComp lies in its reliance on CSI, resulting in high latency and substantial overhead within large-scale IoT networks comprising numerous sensor nodes. Recent research efforts have explored the feasibility of computing the expectation function of the collected data vector without prior knowledge of CSI while maintaining acceptably low estimation errors. The emergence of blind over-the-air computation, which eliminates the need for CSI collection, shows the potential to further reduce computational complexity and latency in IoT networks. In the following paper, Dong et al. propose a novel blind over-the-air computation (BlairComp) framework for low complexity and low latency IoT networks.

Blind Over-the-Air Computation and Data Fusion via Provable Wirtinger Flow

J. Dong, Y. Shi, and Z. Ding, "Blind over-the-air computation and data fusion via provable Wirtinger flow," *IEEE Trans. Signal Process.*, vol. 68, pp. 1136–1151, 2020.

This paper introduces a blind over-the-air computation scheme, BlairComp, designed for low-overhead data fusion in IoT networks. The scheme aims to compute the expectation function of distributed sensing data without the need for channel state information (CSI). The authors propose a randomly initialized Wirtinger flow algorithm to tackle the nonconvex estimation problem inherent in BlairComp. They delve into the population-level state evolution and the dynamics of the approximate state evolution in the blind over-the-air computation problem. The paper introduces the leave-one-out approach and auxiliary sequences to establish the approximate state evolution. The authors establish the global convergence guarantee of Wirtinger flow with random initialization and demonstrate that the estimation error decays exponentially at a linear convergence rate. In the first stage of the algorithm, with a sufficient data sample, the randomly initialized Wirtinger flow iterations can rapidly converge to a locally strong convex and smooth region within a few iterations. The authors also establish that BlairComp achieves highly accurate estimations in this local region. In the second stage of the algorithm, the estimation error experiences exponential decay.

AirComp leverages the superposition property of wireless channels to enable simultaneous transmission from all nodes. However, it does not address the fundamental challenge of high energy consumption, especially in densely deployed IoT networks. This is because the data fusion center must process all concurrently transmitting nodes, leading to an increase in energy consumption as the number of concurrent nodes increases. Therefore, in scenarios with a large number of transmitting nodes, energy efficiency becomes a crucial concern. An intuitive solution lies in reducing the number of nodes engaged in simultaneous transmission while maintaining precise measurements. In the following paper, Basaran et al. explores how to exploit the spatial correlations among sensor measurements to reduce the complexity of the concurrent transmission of all sensor nodes.

Energy-Efficient Over-the-Air Computation Scheme for Densely Deployed IoT Networks

S. T. Basaran, G. K. Kurt, and P. Chatzimisios, "Energy-efficient over-the-air computation scheme for densely deployed IoT networks," *IEEE Trans. Ind. Informat.*, vol. 16, no. 5, pp. 3558–3565, May 2020.

In this paper, Basaran et al. propose an energy-efficient over-the-air computation scheme for densely deployed IoT networks. The proposed scheme leverages the spatial correlations among sensor measurements to reduce energy consumption and enhance estimation accuracy. The authors introduce a minimum mean square error (MMSE) estimation scheme that considers the cross-correlations among sensor measurements. The proposed scheme includes pre-processing and post-processing blocks at each node, where only a subset of nodes transmit their measurements to the fusion center (FC). By exploiting the spatial correlations among sensor measurements, the scheme reduces the complexity of concurrent transmission of all sensor nodes, which significantly increases energy efficiency. The received signal at the FC is modeled as the sum of sensor measurements and additive white Gaussian noise. The goal is to estimate the arithmetic mean of all sensor measurements. The authors derive closed-form expressions for the MMSE estimators and the mean squared error (MSE). Simulation results validate the theoretical expressions and demonstrate that the proposed scheme can nearly double the network lifetime of IoT systems while maintaining acceptable MSE levels.

Over-the-air computation is a crucial component of the "integrated communication and computation" paradigm, where computational capabilities are seamlessly embedded within communication systems. In recent years, another concept known as "integrated sensing and communication" has gained significant attention for integrating sensing capabilities into communication systems. Integrated sensing and communication systems are also responsible for data transmission and computation; however, the computational aspects related to data transmission remain independent and separate. To simplify operations, reduce overhead, and enhance efficiency, there is a natural inclination to merge the functions of sensing, communication, and computation, which has drawn considerable interest among academics. Motivated by this, Li et al. developed a framework called integrated sensing, communication, and computation over-the-air (ISCCO) in the following paper.

Integrated Sensing, Communication, and Computation Over-the-Air: MIMO Beamforming Design

X. Li et al., "Integrated sensing, communication, and computation over-the-air: MIMO beamforming design," *IEEE Trans. Wireless Commun.*, vol. 22, no. 8, pp. 5383–5398, Aug. 2023.

This paper introduces a framework named integrated sensing, communication, and computation over-the-air (ISCCO), designed to support simultaneous data collection, transmission, and computation in IoT applications. The framework integrates sensing, communication, and computation in a multiple-input-multiple-output (MIMO) system. The authors propose the ISCCO framework, which enables concurrent radar sensing and AirComp, and explore two schemes, including both shared and separated schemes, to optimize the tradeoff between sensing accuracy and AirComp accuracy. The separated scheme divides antenna arrays for radar sensing and AirComp, while all antennas transmit a joint waveform for both radar sensing and AirComp in the shared scheme. The paper examines the beamforming designs for both schemes and evaluates their performance using mean squared error (MSE) metrics. The design challenge of MIMO ISCCO lies in the joint optimization of beamformers on both the IoT devices and the server, resulting in a non-convex problem. To address this issue, an algorithmic solution based on the technique of semidefinite relaxation is proposed. The application of ISCCO in target location estimation is also demonstrated. Simulation results indicate that the proposed ISCCO framework outperforms baseline schemes and achieves superior performance in terms of MSE and target location estimation. The authors also highlight future directions and open research problems in this field.

While AirComp typically supports only basic arithmetic operations, such as summation, it has the potential to be extended to include various nomographic functions through appropriate pre-processing and post-processing stages. Over-the-air analog computation allows offloading computation to the wireless environment through carefully designed transmitted signals, thereby reducing the computational burden on local devices. Recent research efforts have explored the use of over-the-air analog computation techniques to emulate specific computational tasks of convolutional neural networks (CNNs). This advancement promises to significantly reduce the computing complexity and energy consumption associated with neural network computations, thereby facilitating the practical deployment of CNNs within IoT networks. In the following paper, Sanchez et al. design and implement the first-of-its-kind convolution that uses over-the-air computation and demonstrates its application for inference tasks in a CNN.

AirNN: Over-the-Air Computation for Neural Networks via Reconfigurable Intelligent Surfaces

S. G. Sanchez et al., "AirNN: Over-the-air computation for neural networks via reconfigurable intelligent surfaces," *IEEE/ACM Trans. Netw.*, early access, Dec. 14, 2022, doi: 10.1109/TNET.2022.3225883.

In this paper, the authors introduce AirNN, an innovative system that performs over-the-air analog computation for convolutional neural networks (CNNs) using reconfigurable intelligent surfaces (RIS). AirNN harnesses the wireless channel to execute convolution operations in the analog domain by utilizing the channel impulse response (CIR) through RIS. The output of a convolution operation is a time series of samples, which can be computed as the sum of the element-wise product between a finite impulse response (FIR) filter and a subset of sequential samples of the input data. The authors draw a parallel between the digital FIR filter in a convolution operation and the CIR of the wireless channel. By programming the RIS to generate specific signal reflections, they were able to emulate different FIR filters and perform convolutions in the analog domain. The RIS network was designed to manipulate the wireless channel's CIR and implement different FIR filters for convolution. By programming the RIS to create the desired channel transformations, the authors achieved the equivalent of digital convolution through over-the-air analog computation. The authors have demonstrated the feasibility of engineering convolution operations with over-the-air computation through a programmable RIS network that is precisely equivalent to its digital counterpart.