Received 24 October 2023; revised 29 November 2023; accepted 22 December 2023. Date of publication 26 December 2023; date of current version 12 January 2024.

Digital Object Identifier 10.1109/OJCOMS.2023.3347354

Joint User Association and Resource Allocation for Hierarchical Federated Learning Based on Games in Satisfaction Form

PANAGIOTIS CHARATSARIS, MARIA DIAMANTI (Member, IEEE), AND SYMEON PAPAVASSILIOU (Senior Member, IEEE)

Institute of Communication and Computer Systems, School of Electrical and Computer Engineering, National Technical University of Athens, 15780 Zografou, Greece

CORRESPONDING AUTHOR: S. PAPAVASSILIOU (e-mail: papavass@mail.ntua.gr)

This work was supported in part by the European Commission through the Horizon Europe/JU SNS Project Hexa-X-II under Grant 101095759.

ABSTRACT Hierarchical Federated Learning (HFL) has emerged to overcome the shortcomings of conventional Federated Learning (FL) due to communication obstacles between the end users and the cloud server and the congestion at the backhaul of wireless network implementations. In this paper, we consider a wireless user-edge-cloud HFL network where the transmissions of the users' local model parameters to the edge are multiplexed via the Non-Orthogonal Multiple Access (NOMA) technique. The joint problem of association and uplink transmission power allocation of the users to the edge is formulated and solved as a non-cooperative game in satisfaction form. Diverging from the prevailing research that proposes centralized solution concepts, each user makes autonomous decisions regarding its association and power level so as to attain a minimum acceptable tradeoff of three vital network factors. The latter includes the global model's training accuracy and the users' consumed energy and time during transmission. Different types of equilibria are explored, i.e., the Satisfaction Equilibrium (SE) and Minimum Efficient Satisfaction Equilibrium (MESE) which not only fulfills users' minimum tradeoff but also minimizes the overall network's cost. Algorithms based on Reinforcement Learning (RL) and Best Response Dynamics (BRD) are, then, devised to conclude the SE and MESE points. The proposed framework is evaluated via modeling and simulation, verifying its efficiency in achieving an equitable balance in the network.

INDEX TERMS Hierarchical federated learning, game theory, games in satisfaction form, user association, power allocation.

I. INTRODUCTION

THE PROLIFERATION of the Internet of Things (IoT) and social networking has provoked individuals to generate massive amounts of data through interconnected devices. This offers great potential for leveraging Machine Learning (ML) models in various applications, such as natural language processing, image processing, and 6G wireless communications, just to name a few of them. However, centralized ML faces challenges due to the increase in the complexity of the aforementioned applications and the growing concerns about data privacy. In this context, Federated Learning (FL) offers a viable solution by enabling collaborative training of ML models across distributed end-user devices while preserving data privacy [1]. The end users collectively train a global ML model based on their local data by solely exchanging their local model updates with a remote server. From a communication perspective, though, several challenges arise related to increased backhaul network traffic or even connectivity issues that hinder the performance of distributed learning when considering implementations of FL over wireless networks [2].

Hierarchical Federated Learning (HFL), in turn, introduces the notion of hierarchy in the FL procedure by empowering devices at different levels across the computing continuum to collaborate effectively with each other [3]. In particular, HFL suggests adding an extra layer of edge model aggregation where edge servers facilitate the aggregation and transmission of end users' model parameters to a remote server in the cloud. In this way, better resource utilization across the continuum can be achieved while reducing backhaul network traffic and communication costs in both the convergence time of the FL procedure and consumed energy at the end-user devices. Nevertheless, an inconvenient assignment of users to the different edge servers can cause an unequal distribution of data among the servers and an imbalanced traffic load on the Radio Access Network (RAN), bringing exactly opposite effects in the HFL procedure. To reap the maximum benefits of such a hierarchical structure, several research works, e.g., [4], [5], [6], [7], [8], are devoted to the appropriate user association to the available edge servers and the allocation of the wireless resources.

In this paper, we strive to address this particular problem in a wireless HFL network, where the users' transmissions of their local model parameters to the edge are multiplexed using the Non-Orthogonal Multiple Access (NOMA) technique. Different from [4], [5], [6], [7], [8], we focus on the socio-economic perspective of wireless HFL networks. We specifically target the tradeoff between the global model's training accuracy and users' incurred communications energy and time overhead, serving as an incentive for the users to invest their resources and participate in the HFL procedure. Each user pursues its individual minimum acceptable tradeoff by autonomously determining its association and uplink transmission power to the edge, while the interactions and interdependent decisions among the different users are modeled as a game in satisfaction form [9], [10]. Contrary to the conventional normal-form non-cooperative games that aim at the players' utility maximization, this form of games targets the satisfaction of the players' minimum acceptable Quality of Service (QoS) requirements, which are mapped to the users' minimum acceptable accuracy-time-energy tradeoff value [11].

Overall, in this paper, the joint problem of user association and uplink transmission power allocation in a wireless HFL network employing power-domain NOMA is addressed as a game in satisfaction form to achieve a tradeoff among accuracy, time, and energy overheads for the individual users. Different types of equilibria are scrutinized therein, such as the Satisfaction Equilibria (SE) and Minimum Efficient Satisfaction Equilirbia (MESE) that not only satisfy the users' minimum desired tradeoff but also result in the overall network's minimum cost. Appropriate algorithms are also designed for the derivation of the corresponding equilibria.

A. RELATED WORK

Numerous studies have delved into FL's application and demonstrated its effectiveness under various settings so far [12]. Especially for implementations over wireless networks, HFL is currently a vibrant field of investigation due to its capability to address several hurdles encountered in traditional FL. A particular implementation of HFL considers the organization of the different users in the network into clusters and the determination of a corresponding cluster-head to serve as a mediator to the rest of the users' transmissions of their local model parameters to the cloud [13]. Other HFL implementations are founded on the typical user-edge-cloud structure, e.g., [14], [15]. Specifically, the work in [14] constitutes one of the first in the field to provide convergence guarantees for the HFL procedure over wireless networks. Going one step further and accounting for user mobility, the work in [15] demonstrates how the latter can be leveraged to enhance learning performance by properly selecting the users that participate in the HFL procedure at each learning round based on their position. This allows for the application of HFL (and FL in general) in the context of the Internet of Vehicles (IoV) [16] for collaborative training of an ML model while keeping data locally, while more elaborate applications of HFL can be used hereupon to combat malicious attacks [17].

The performance of wireless FL is heavily contingent upon the quality of the communication between the various network entities and the allocation of radio resources. Considering a single-server multi-user FL system, to begin with, several research works exist tackling the joint user selection/scheduling and resource allocation problem under different network settings and optimization objectives, e.g., [18], [19], [20], [21], [22]. Both [18] and [19] consider the implementation of FL over Orthogonal Frequency Division Multiple Access (OFDMA) networks and optimize the user selection, uplink transmission power, bandwidth, and CPU frequency allocation at the user devices. In [18], the minimization of the consumed energy is pursued, while [19] accounts additionally for the FL procedure's convergence time minimization. Nevertheless, Orthogonal Multiple Access (OMA) techniques impose significant limitations regarding the number of concurrently supported users over the same time, frequency, or code-domain resources, reducing the network's spectral efficiency. To combat this issue, NOMA has also been considered in FL networks, e.g., [20], [21], allowing multi-user multiplexing by employing advanced interference cancellation techniques at the receiver. In [20] and [21] a similar variable set with [18], [19] is optimized, focusing on the sum-rate maximization and energy consumption minimization, respectively. Aiming at the global model's accuracy enhancement, the maximization of the number of users utilized in the FL procedure is targeted in [22] by optimally controlling the transmission power, network bandwidth, and user selection.

In HFL structures, instead, the joint problem of association and resource allocation between the user and edge tiers is mainly studied in the literature to manage the imbalanced user and data distribution across the RAN, e.g., [4], [5], [6], [7], [8]. In [4], an analysis of the impact of user association in HFL networks with Independent Identically Distributed (IID) and non-IID data distributions to the different edge servers is performed, which is extended in [5] by concurrently accounting for the effect of resource allocation in terms of bandwidth. The joint problem of user association and resource allocation is, then, solved in [5] to minimize the learning latency of the FL procedure while a similar problem is addressed in [6] to minimize the weighted sum of energy and time overheads across an HFL network. Contrary to the aforementioned works that consider OFDMA-based environments, the authors in [7] seek the minimization of both latency and energy consumption and the maximization of the learning model's accuracy under a NOMA implementation by optimizing the users' uplink transmission power along with their association. A centralized solution is proposed there which, however, is difficult to scale as the size of the network increases, given that the optimization falls in the Mixed Integer Non-Linear Programming (MINLP) NP-hard problems. Distributed optimization frameworks based on Game Theory provide effective alternatives to this scalability issue. The work in [8] provides a particular example where Stackelberg and evolutionary non-cooperative games are used in HFL networks to give solutions to optimization problems.

However, what has been significantly neglected so far is that the users engaged in HFL aim at high model accuracy without necessarily excessively consuming their power resources. The latter describes a tradeoff that needs to be satisfied, different from maximizing the global model's accuracy subject to the users' energy and time constraints. In this context, the theory of non-cooperative games in satisfaction form [9] allows us to achieve a minimum acceptable tradeoff between model accuracy and consumed energy and time for communication in a distributed manner. The users can self-organize by determining their association and uplink transmission power as players in a non-cooperative game while pursuing the satisfaction of a tradeoff value instead of its maximization. Concurrently, the benefits of a game-theoretic solution are inherited by distributing decision-making across the users related to reduced computational complexity, enhanced user privacy, and dynamic adaptation to changing conditions. Especially under this form of games, different types of equilibria can be achieved that consider either the exclusive satisfaction of the involved users (i.e., SE [23]) or even the simultaneous minimization of the cost that is incurred to the overall network (i.e., MESE [24]). A wide range of applications of the games in satisfaction form exist, including user association in heterogeneous networks [25], channel allocation [26], and minimization of base stations' energy consumption [27].

B. CONTRIBUTIONS & OUTLINE

In this paper, we introduce and solve – for the first time in the literature of wireless HFL networks – the joint problem of association and uplink transmission power allocation between the users and the edge while targeting to achieve a minimum acceptable accuracy-time-energy tradeoff from the users' side in a distributed way. The pursued tradeoff regards the produced global model's accuracy and the incurred energy and time consumption to the users due to communication with the edge. In this context, the main contributions of this paper are summarized as follows.

1) The joint problem of association and uplink transmission power allocation of the users to the edge in a wireless HFL network using power-domain NOMA, is modeled as a game in satisfaction form, such that each user pursues its accuracy-time-energy tradeoff autonomously.
2) Different types of equilibria, i.e., SE and MESE, are analyzed and discussed in terms of their existence and uniqueness, concluding with different solution outcomes that range from the users' minimum acceptable tradeoff satisfaction to the overall network's optimal performance in terms of minimum incurred cost.
3) Distributed algorithms are introduced that are executed autonomously by the users in the network to conclude the different equilibria based on Reinforcement Learning (RL) [9] and Best Response Dynamics (BRD) [28].
4) Numerical evaluation based on modeling and simulation demonstrates the effectiveness of the proposed framework in achieving the desired global model's accuracy and users' incurred energy and time overhead tradeoff, successfully fulfilling the HFL procedure's learning objective while securing the end users' beneficial participation.

The remainder of the paper is organized as follows. Section II presents the modeling of the HFL procedure, user interference, and accuracy-time-energy tradeoff. In Section III, the formulation of the game in satisfaction form is provided and the different equilibria are analyzed. Section IV presents the devised algorithms to derive the different equilibria and thus, solution outcomes. Section V regards the framework's performance evaluation and Section VI concludes the paper.

II. SYSTEM MODEL

We consider a wireless HFL network, as illustrated in Fig. 1, where multiple edge servers and users collaboratively train a global learning model, coordinated by a model owner. We denote the set of edge servers as $\mathcal{M} = \{1, \dots, m, \dots, M\}$ and the set of users as $\mathcal{N} = \{1, \dots, n, \dots, N\}$. The edge servers, participating in the training phase, are hosted and collocated with the RAN's base stations, whereas the model owner resides in a cloud server. Each edge server $m \in \mathcal{M}$ assembles the local models of its connected users and produces an intermediate edge model, while the same procedure is followed by the cloud server that gathers and aggregates the edge models to create the global model. We suppose that the serving areas of the different edge servers are overlapping, such that each user autonomously selects the one edge server to associate with and broadcast its local model parameters. The set of users that are associated with each edge server $m \in \mathcal{M}$ is indicated by $\mathcal{N}_m =$

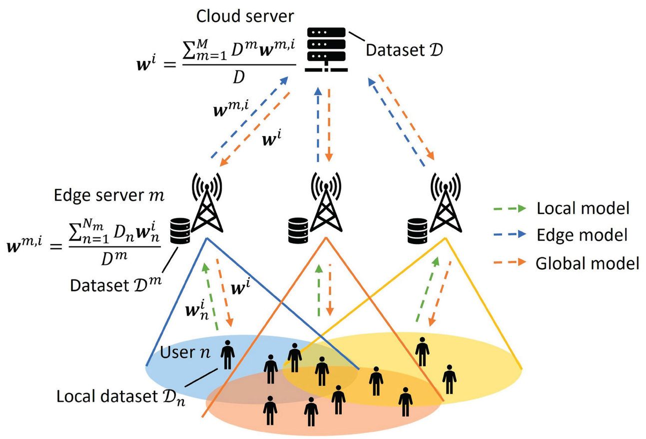

FIGURE 1. Overview of the wireless hierarchical federated learning (HFL) network.

$\{1, \ldots, n, \ldots, N_m\}$. Moreover, we define variable $a_{n,m} \in \{0, 1\}$ to capture the relationship between a user and an edge server, such that a user n is associated with edge server m if $a_{n,m} = 1$, and vice versa.

To perform the learning task, each user n possesses a set of training data denoted by $\mathcal{D}_n = \{\mathbf{x}_j, y_j\}_{j=1}^{D_n}$ of cardinality D_n. Considering that the HFL procedure regards a classification learning task, $\mathbf{x}_j$ is the j-th input sample and y_j is the corresponding class. Overall, we consider that there exist C different sample classes indicated by the set $\mathcal{C} = \{1, \ldots, c, \ldots, C\}$. The data volume resulting from the users $n \in \mathcal{N}_m$ that are associated with edge server m is represented as $\mathcal{D}^m = \bigcup_{\forall n \in \mathcal{N}_m} \mathcal{D}_n = \{\mathbf{x}_j, y_j\}_{j=1}^{D^m}$. The total dataset of the network is $\mathcal{D} = \bigcup_{\forall n \in \mathcal{N}} \mathcal{D}_n = \{\mathbf{x}_j, y_j\}_{j=1}^{D}$. Table 1 lists the key notation of the paper.

A. HIERARCHICAL FEDERATED LEARNING MODEL

The considered HFL process comprises three sequential and iterative phases, as shown in Fig. 1, that regard the local model training and the edge and cloud model aggregations. In more detail, the operation of the HFL procedure is as follows. In the local model training phase, the users practically aim to minimize the loss of their local model that is expressed by the function:

$$F_n(\mathcal{D}_n, \mathbf{w}_n) = \frac{1}{D_n} \sum_{j=1}^{D_n} f_n\big(\mathbf{x}_j, y_j, \mathbf{w}_n\big). \tag{1}$$

where f_n is the empirical loss function of the j-th data sample of the local dataset D_n, expressing the classification error of the model for this sample. In this paper, we adopt the cross-entropy loss function [29]. Toward minimizing the aforementioned loss, a number of κ_1 local model updates/iterations is performed, where each user's n local model parameters $\mathbf{w}_n$ are updated via the typical gradient descent rule:

$$\mathbf{w}_n^i = \mathbf{w}_n^{i-1} - \eta \nabla F_n\Big(D_n, \mathbf{w}_n^{i-1}\Big), \tag{2}$$

TABLE 1. Table of key notation.

Notation	Definition
$\mathcal{M}$	Set of edge servers
$\mathcal{N}$	Set of users
$\mathcal{N}_m$	Set of users associated with server m
$\mathcal{D}_n$, $\mathcal{D}^m$, $\mathcal{D}$	Datasets of user n, users in $\mathcal{N}_m$, and total network
$\{\mathbf{x}_j, y_j\}$	j-th input sample and corresponding class of dataset $\mathcal{D}$
$\mathcal{C}$	Set of dataset's $\mathcal{D}$ classes
P_m^c	Percentage of samples of class c in D^m
H_m	Information entropy of edge server m
F_n	Mean loss of user's n samples
f_n	Loss of user's n individual samples
η	Training learning rate
κ_1, κ_2, K	Local, edge servers', and global model iterations
$\mathbf{w}_n^i$	Local model parameters of user n at iteration i
$Z(\mathbf{w}_n)$	Data size of user's n local model parameters
W_m, W	Bandwidth of edge server m and total network
$G_{n,m}$	Channel gain between user n and edge server m
$a_{n,m}$	Association indicator between server m and user n
p_n	Uplink transmission power of user n
p_n^{max}	Maximum transmission power of user n
$R_{n,m}$	Uplink throughput of user n to edge server m
$T_{n,m}^{tx}$	Transmission time of user n to edge server m
$E_{n,m}^{tx}$	Transmission energy of user n to edge server m
w_e, w_t	Energy, Time weight factors
u_n	Utility function of user n
u_n^{thr}	Minimum acceptable tradeoff of user n
$\mathcal{S}$	Strategy space of user n
f_n	Requirements set of user n
c_n	Cost function of user n

with i denoting the iteration index and η being the training's learning rate.

In the second edge model aggregation phase, each server m gathers the local parameters $\mathbf{w}_n^i$ of its associated users $n \in \mathcal{N}_m$ and derives an updated edge model $\mathbf{w}^{m,i}$ by employing the following model aggregation rule [30]:

$$\mathbf{w}^{m,i} = \frac{\sum_{n=1}^{N_m} D_n \mathbf{w}_n^i}{D^m}, \quad \text{if} \quad i \mod \kappa_1 = 0. \tag{3}$$

The updated edge model is then transmitted back to the users to replace their existing local model. This second phase is repeated for κ_2 iterations until the overall HFL procedure is concluded.

Last, similar to the edge model aggregation phase, the model owner, i.e., cloud server, aggregates the intermediate learning models of the edge servers and produces the global model parameters $\mathbf{w}^i$ [30]:

$$\mathbf{w}^{i} = \frac{\sum_{m=1}^{M} D^m \mathbf{w}^{m,i}}{D}, \quad \text{if} \quad i \mod \kappa_1\kappa_2 = 0. \tag{4}$$

The updated global model is fed back to the users to facilitate the next HFL iteration. The overall HFL process is repeated for K iterations until a desired accuracy level is reached.

B. WIRELESS COMMUNICATION MODEL

The transmissions performed in the wireless access part of the network, i.e., between the users and the edge servers, are facilitated over a total bandwidth of W [Hz]. The total bandwidth is further divided equally and allocated to the different base stations and their corresponding edge servers, such that $W_m = \frac{W}{M}$ [Hz] is the available bandwidth for communication with edge server m. Subsequently, the different users' $n \in \mathcal{N}_m$ transmissions to edge server m are multiplexed via the power-domain Non-Orthogonal Multiple Access (NOMA) technique.

We denote as $G_{n,m}$ the channel gain between user $n \in \mathcal{N}_m$ and edge server m, defined as $G_{n,m} = \rho d_{n,m}^{-\alpha}$, where ρ [dB] is the path loss at the reference distance of 1 m, $d_{n,m}$ [m] is the Euclidean distance between user n and edge server m, and α is the path loss exponent. Without loss of generality, we assume that the channel gains between users $n \in \mathcal{N}_m$ and edge server m are sorted in ascending manner, $G_{1,m} \leq \cdots \leq G_{n,m} \leq \cdots \leq G_{N_m,m}$, and the decoding of the signals begins from the highest channel gain user using the Successive Interference Cancellation (SIC) technique. The achieved uplink throughput of user $n \in \mathcal{N}_m$ to edge server m is calculated as:

$$R_{n,m} = W_m \log_2\left(1 + \frac{p_n G_{n,m}}{\sum_{n'=1}^{n-1} a_{n',m} p_{n'} G_{n',m} + I_0^m}\right), \quad (5)$$

where p_n [W] is the corresponding uplink transmission power of user n and I_0^m [dBm/Hz] is the power spectral density of zero-mean Additive White Gaussian Noise (AWGN). Well-aligned with relative works that focus on the resource allocation aspect of the considered problem, e.g., [7], [8], we assume perfect Channel State Information (CSI) at both edge server and user sides to perform SIC and distributed decision-making, respectively. Specifically, we assume that the CSI is estimated at the edge server by the training sequences transmitted by the users that are a priori known to the edge server, and this information is also reported to the user.

Given the communication model between the users and the edge servers, the time overhead of user n for transmitting its local model parameters to its associated edge server m is given by [31]:

$$T_{n,m}^{tx} = \frac{Z(\mathbf{w}_n)}{R_{n,m}} \quad [s]. \quad (6)$$

where $Z(\mathbf{w}_n)$ is the data size in bits of its local model parameters' vector $\mathbf{w}_n$ transmitted to update the edge model. Furthermore, the corresponding incurred energy overhead is:

$$E_{n,m}^{tx} = \frac{Z(\mathbf{w}_n) p_n}{R_{n,m}} \quad [J]. \quad (7)$$

It should be noted that the modeling of the communications performed at the backhaul network between the edge servers and the model owner, i.e., the cloud, is beyond the scope of this paper. Although the latency in the backhaul due to interference among the transmissions and limited bandwidth can affect the speed of model updates at the cloud server, the emphasis of this paper is placed on the wireless access network part. Joint uplink transmission power control and bandwidth allocation at the backhaul network are required, which can follow the related literature in Integrated Access and Backhaul (IAB) networks [32], [33], [34], [35]. The joint optimization of access and backhaul network communications in wireless HFL networks is part of our future work.

C. ACCURACY-TIME-ENERGY TRADEOFF MODEL

In this section, we define each end user's n utility function u_n that serves as a measure of the achieved tradeoff between the global model's training accuracy, and the incurred time and energy overheads at the users' side. Precisely, a user's n utility is given by:

$$u_n = \sum_{m=1}^{M} a_{n,m} \frac{\sum_{m'=1}^{M} H_{m'}(\mathcal{D}^m)}{w_e E_{n,m}^{tx} + w_t T_{n,m}^{tx}}. \quad (8)$$

The numerator of (8) captures the users' data distribution among the different edge servers quantified by the information entropy metric denoted as $H_m(\mathcal{D}^m), \forall m$, while the denominator represents the weighted consumed time and energy for transmitting the local model parameters from the users to the edge, with $w_e, w_t \in [0, 1]$ are appropriate weight factors, such that $w_e + w_t = 1$.

Especially with reference to the numerator, the information entropy $H_m(\mathcal{D}^m)$ at edge server m that possess data volume $\mathcal{D}^m$ is [36]:

$$H_m(\mathcal{D}^m) = -\sum_{c=1}^{C} P_m^c(\mathcal{D}^m) \cdot \log(P_m^c(\mathcal{D}^m)), \quad (9)$$

where $P_m^c(\mathcal{D}^m)$ is the percentage of samples of class c in $\mathcal{D}^m$. Apparently, high values of the information entropy metric indicate a network configuration that closely resembles the IID case. In that case, all edge servers are associated with users whose aggregated data set includes samples from all present classes in the classification problem. The latter ideal scenario enhances the training accuracy in the HFL network [4], implying a direct relation between the value of the information entropy and the achieved training model accuracy. Capitalizing on this relation allows us to indirectly account for the HFL model's accuracy in the users' utility function while disentangling the users' optimization problem from a variable directly tied to the learning process. In this way, the proposed joint user association and resource allocation framework is executed independently and before the HFL procedure, such that the complexity and convergence behavior of the HFL procedure remain unaffected, following the related analyses in the literature [14].

As a concluding remark, by selecting convenient actions concerning its association and uplink transmission power, each end user can guarantee a utility threshold u_n^{thr}, i.e., $u_n \geq u_n^{thr}$, targeting a desirably high network's entropy and low time and energy consumption for the whole learning process.

III. JOINT USER ASSOCIATION & RESOURCE ALLOCATION BASED ON GAMES IN SATISFACTION FORM

In this section, the joint problem of user association and uplink transmission power allocation in the access part of the

HFL network is studied to accommodate the transmission of the users' local model parameters to the edge. This joint problem is a MINLP NP-hard problem, and thus, deriving an optimal solution within polynomial time is practically infeasible. For this reason, we propose an approach based on a non-cooperative game in satisfaction form and examine different types of equilibria that satisfy the minimum acceptable accuracy-energy-time tradeoff for each user in the HFL network. The concluded equilibria dictate the selected associations and transmission powers of the users that are not necessarily optimal in strict terms but constitute stable solution outcomes, from which no user has the motivation to deviate.

A. GAME IN SATISFACTION FORM

The non-cooperative game in satisfaction form between the users is described by the following tuple:

$$\mathcal{G} = \left(\mathcal{N}, \{\mathcal{S}_n\}_{n\in\mathcal{N}}, \{f_n\}_{n\in\mathcal{N}}\right), \tag{10}$$

consisting of three components.

1) The set of players $\mathcal{N}$, i.e., the users participating in the learning process.
2) The space $\mathcal{S}_n = \{s_n = (p_n, x_n) | p_n \in [0, p_{max}], x_n \in \mathcal{M}\}$ of user's n actions of cardinality S_n, where $x_n \in \mathcal{M}$ is an integer variable indicating the edge server $m \in \mathcal{M}$ that user n chooses to associate with, and p_{max} [W] is the maximum feasible power level of each user. The action space $\mathcal{S}_n$ includes all possible combinations s_n of transmission powers and edge server selections. Given the edge server selection parameter x_n, the corresponding binary association indicators $a_{n,m}$ can be derived by backward induction as:
$$a_{n,m} = \begin{cases} 1, & \text{if } m = x_n, \\ 0, & \text{otherwise.} \end{cases} \tag{11}$$
3) The requirements set $f_n(\mathbf{s}_{-n}) : \mathcal{S}_{-n} \longrightarrow 2^{S_n}$ which is defined as:
$$f_n(\mathbf{s}_{-n}) = \left\{ s_n \in \mathcal{S}_n | u_n(s_n, \mathbf{s}_{-n}) \geq u_n^{thr} \right\}, \tag{12}$$
where $\mathbf{s}_{-n}$ denotes the actions of all users $n \in \mathcal{N}$ except for user n. In particular, the set f_n consists of all the actions of user n that guarantee the satisfaction of its minimum acceptable tradeoff u_n^{thr}, supposing having knowledge of the actions selected by the rest of the users.

Given that the users are not aiming at the maximization of their utility function under the considered game structure, the typical Nash Equilibrium (NE) [37] derived as an outcome from the Best Response Dynamics algorithm [37] is not the solution we are seeking for. In the sequel, different types of satisfaction equilibria – as they are called – are scrutinized, yielding different, though suitable, solutions to our problem.

B. SATISFACTION EQUILIBRIUM (SE)

The general outcome of a game in satisfaction form that takes into consideration the users' satisfaction is called Satisfaction Equilibrium (SE). The definition and existence of this type of equilibrium are provided in the following proposition.

Definition 1 (Satisfaction Equilibrium and Existence): An action profile $\mathbf{s}^* = (s_1^*, \ldots, s_n^*, \ldots, s_N^*)$ is an SE for game $\mathcal{G}$ if:

$$s_n^* \in f_n(\mathbf{s}_{-n}), \forall n \in \mathcal{N}. \tag{13}$$

Therefore, based on Kakutani's fixed point theorem [38], game $\mathcal{G}$ has at least one SE if the action spaces $\{\mathcal{S}_n\}_{n\in\mathcal{N}}$ are non-empty, convex, and compact sets, and the requirements set $F(\mathbf{s}) = \{f_1(\mathbf{s}_{-1}), \ldots, f_N(\mathbf{s}_{-N})\}$ of all players is a non-empty, convex set of their actions $\{\mathcal{S}_n\}_{n\in\mathcal{N}}$, and has a closed graph.

Apparently, when an SE is achieved, all users' minimum acceptable tradeoff is satisfied and none of them has the incentive to change its action. The existence of at least one SE for the examined game $\mathcal{G}$ depends exclusively on the requirements set of each user and is irrelevant to the form or the properties of the utility function. Consequently, to find the necessary conditions for the existence of this equilibrium, we can extend an existing fixed point theorem.

Definition 1 implies the necessity for feasible system initialization to guarantee the satisfaction of all users. The existence, however, of an SE point cannot guarantee its uniqueness. Specifically, it is challenging to define different subsets of the requirements set $\{f_n\}_{n\in\mathcal{N}}$ that can imply a unique combination of actions of the users that keeps them satisfied.

C. EFFICIENT SATISFACTION EQUILIBRIUM (ESE)

In the general case discussed above, the SE of the problem under investigation is not unique, meaning that there exist multiple combinations of user actions that keep them satisfied. Nevertheless, these combinations can be distinguished from each other with respect to the cost incurred to the users.

In our analysis, we consider that a user's cost is a function of the power level required to transmit its local model parameters to the edge, i.e., $c_n(s_n) = c_n(p_n)$, taking indirectly into account its particular association with respect to its actual distance from the selected edge server, their in-between wireless channel conditions, and the interference sensed from the rest of the users communicating with the same server. Obviously, it is more beneficial for the users to satisfy their minimum acceptable tradeoff with the minimum possible cost at the same time. As a logical consequence of the above observation, in the sequel, we define the Efficient Satisfaction Equilibrium (ESE) point.

Definition 2 (Efficient Satisfaction Equilibrium): An action profile $\mathbf{s}^* = (s_1^*, \ldots, s_n^*, \ldots, s_N^*)$ is an ESE for game $\mathcal{G}$, with cost function $\{c_n\}_{n\in\mathcal{N}}$, if the following two conditions concurrently hold true:

$$s_n^* \in f_n(\mathbf{s}_{-n}), \forall n \in \mathcal{N}, \tag{14a}$$

$$c_n(s_n) \geq c_n\left(s_n^*\right), \forall n \in \mathcal{N}, \forall s_n \in f_n\left(\mathbf{s}_{-n}^*\right). \tag{14b}$$

Given that at least one SE exists for the studied game, as declared in Definition 1, it suffices to confirm that there exists an SE where the users satisfy their minimum acceptable tradeoff with the minimum possible individual cost. Under this prism, we define the best response function $\mathcal{BR}_n(\mathbf{s}_{-n})$ of each user n that yields the optimal action s_n, given the actions of the rest users, as: $\mathcal{BR}_n(\mathbf{s}_{-n}) = \{s_n = (p_n, x_n) \in \mathcal{S}_n{:}s_n = \arg\min_{s_n \in f_n(\mathbf{s}_{-n})} c_n(p_n)$. Alternatively, the optimization problem that each user n seeks to solve to determine an ESE is:

$$\begin{aligned} (\mathbf{P}) : \min_{(p_n, x_n)} \quad & p_n \\ \text{s.t.} \quad & p_n \in \left[0, p_{max}\right], x_n \in \mathcal{M}, \\ & u_n(s_n, \mathbf{s}_{-n}) \geq u_n^{thr}. \end{aligned} \tag{15}$$

The existence of a solution for problem (15), i.e., an ESE for the game $\mathcal{G}$, is guaranteed under the case that the requirements set $f_n, \forall n \in \mathcal{N}$ is non-empty, as stated in Proposition 1. To facilitate the analysis regarding the existence of at least one ESE, subsequently, we summarize the fixed point theorem of Knaster and Tarksi [39].

Theorem 1 (Tarski and Knaster's Fixed Point Theorem): Let $\mathcal{T}$ be a complete lattice and let $h : \mathcal{T} \longrightarrow \mathcal{T}$ be a monotonic function. Then, the set of fixed points of h in $\mathcal{T}$ is also a complete lattice.

Proposition 1 (Existence of ESE): If the non-cooperative game $\mathcal{G}$ in satisfaction form with cost function $\{c_n\}_{n\in\mathcal{N}}$ and utility function $\{u_n\}_{n\in\mathcal{N}}$ has non-empty requirements set $f_n, \forall n \in \mathcal{N}$, then it possesses at least one ESE.

Proof: We define a lattice $\mathcal{T} = <\mathcal{S}, \leq>$, where $\mathcal{S} = \cup_{n\in\mathcal{N}}\mathcal{S}_n$ is the overall action space of game $\mathcal{G}$ and $\leq$ represents the component-wise less or equal operation. All subsets of the lattice have both a supremum and an infimum, as the transmission power of the users lies within an interval of a closed set $[0, p_{max}]$, and thus, $\mathcal{T}$ is a complete lattice. Concerning the monotonic function $h : \mathcal{T} \longrightarrow \mathcal{T}$, it is defined as follows:

$$h(\mathbf{s}) = (\mathcal{BR}_1(\mathbf{s}_{-1}), \ldots, \mathcal{BR}_N(\mathbf{s}_{-N})), \forall \mathbf{s} = (s_1, \ldots, s_N) \in \mathcal{S}.$$

Each user's n utility is a quasiconcave function with respect to its power p_n, i.e., strictly increasing on p_n [35]. Accordingly, in case all competing users change their actions to alternatives with higher or equal transmission power, user n will either still satisfy its minimum acceptable tradeoff or will have to increase its power. This can be formally written as: $\forall \mathbf{s}, \mathbf{s}' \in \mathcal{S}{:}p_n^{(\mathbf{s})} \leq p_n^{(\mathbf{s}')} \Rightarrow (\mathcal{BR}_1(\mathbf{s}_{-1}), \ldots, \mathcal{BR}_N(\mathbf{s}_{-N})) \leq (\mathcal{BR}_1(\mathbf{s}'_{-1}), \ldots, \mathcal{BR}_N(\mathbf{s}'_{-N})) \iff h(\mathbf{s}) \leq h(\mathbf{s}')$, where $p_n^{(\mathbf{s})}$ indicates the user's n uplink transmission power as determined based on action profile $\mathbf{s}$. This proves that h is an order-preserving function, and based on Tarski-Kraskel's theorem a fixed point of function h exists:

$$\begin{aligned} & \exists\, \mathbf{s} \in \mathcal{S} : \mathbf{s} = h(\mathbf{s}) \\ & \iff (s_1, \ldots, s_N) = (\mathcal{BR}_1(\mathbf{s}_{-1}), \ldots, \mathcal{BR}_N(\mathbf{s}_{-N})). \end{aligned}$$

This means that all users have played their best response to the rest of the users' actions, and therefore $\mathbf{s}$ is an ESE for game $\mathcal{G}$. ■

D. MINIMUM EFFICIENT SATISFACTION EQUILIBRIUM (MESE)

The equilibria studied so far address the users' satisfaction, even with minimum cost for them when referring to the ESE, but provide no guarantees for the total cost that is incurred to the network. The equilibrium point that is complementary to the above and results in the minimum total cost from the overall network's perspective is called Minimum Efficient Satisfaction Equilibrium (MESE). It should be noted that the term "minimum network cost" refers to the summation of the individual users' costs.

Definition 3 (Minimum Efficient Satisfaction Equilibrium): An action profile $\mathbf{s}^* = (s_1^*, \ldots, s_n^*, \ldots, s_N^*)$ is a MESE for game $\mathcal{G}$, with cost function $\{c_n\}_{n\in\mathcal{N}}$ and set of action profiles $\{\mathrm{E}\}$ that are ESEs if the following conditions concurrently hold true:

$$s_n^* \in f_n(\mathbf{s}_{-n}), \forall n \in \mathcal{N}, \tag{16a}$$

$$c_n(s_n) \geq c_n\left(s_n^*\right), \forall n \in \mathcal{N}, \forall s_n \in f_n\left(\mathbf{s}_{-n}^*\right), \tag{16b}$$

$$\sum_{n\in\mathcal{N}} c_n(e_n) \geq \sum_{n\in\mathcal{N}} c_n\left(s_n^*\right), \forall \mathbf{e} \in E. \tag{16c}$$

The existence of a MESE is closely related to that of an SE and is discussed in the following proposition.

Proposition 2 (Existence of MESE): If the non-cooperative game $\mathcal{G}$ in satisfaction form with cost function $\{c_n\}_{n\in\mathcal{N}}$ and utility function $\{u_n\}_{n\in\mathcal{N}}$ has $f_n(\cdot) \neq \emptyset, \forall n \in \mathcal{N}$ for every input, then it possesses at least one MESE.

Proof: Based on Proposition 1, if a game $\mathcal{G}$ possesses a combination of actions that satisfy the constraints $f_n, \forall n \in \mathcal{N}$, then the game admits at least one ESE. As a logical consequence, one combination of them yields the minimum cost for the network, implying the existence of at least one MESE point. ■

A particular feature of the MESE point is that upon its existence, the obtained MESE point is also unique.

Proposition 3 (Uniqueness of MESE): If the non-cooperative game $\mathcal{G}$ has a MESE point $\mathbf{s}^*$, that point is unique considering a given association between the users and the edge servers.

Proof: Let $\{E\}$ denote the set of all action profiles in the action space $\mathcal{S}$ that are ESEs for game $\mathcal{G}$. The proposition is proved by contradiction.

Consider that two different MESE points $\hat{\mathbf{s}}$ and $\bar{\mathbf{s}}$ exist for game $\mathcal{G}$, such that $\exists\, n \in \mathcal{N}, c_n(\hat{s}_n) \neq c_n(\bar{s}_n)$ and the points are not identical profiles regarding the power allocation. Note that the uniqueness of the MESE point is proved upon a given association that is also considered to be identical for the two MESE points $\hat{\mathbf{s}}$ and $\bar{\mathbf{s}}$. According to Definition 3, the two MESE points should offer exactly the same network cost, which will be lower than that offered by any other ESE denoted as $\mathbf{s}^{ESE}$, i.e.,

$$\sum_{\forall n\in\mathcal{N}} c_n(\hat{s}_n) = \sum_{\forall n\in\mathcal{N}} c_n(\bar{s}_n) \leq \sum_{\forall n\in\mathcal{N}} c_n\left(s_n^{ESE}\right), \forall s \in \{E\}.$$

Furthermore, since the two points are not identical and the cost function depends exclusively on the uplink transmission power, it holds that $\exists\, n_1 \in \mathcal{N}, p_{n_1}^{(\hat{\mathbf{s}})} \neq p_{n_1}^{(\bar{\mathbf{s}})}$, for which, without loss of generality, we assume that $p_{n_1}^{(\hat{\mathbf{s}})} \leq p_{n_1}^{(\bar{\mathbf{s}})}$. Given that the two MESE points yield the same network cost, $\exists\, n_2 \in \mathcal{N}, p_{n_2}^{(\hat{\mathbf{s}})} \geq p_{n_2}^{(\bar{\mathbf{s}})}$, such that the summation of all users' cost leaves the total network cost over the two MESE points unchanged, i.e., $\sum_{\forall n\in\mathcal{N}} c_n(\hat{s}_n) = \sum_{\forall n\in\mathcal{N}} c_n(\bar{s}_n)$.

Let now choose randomly an action profile $\mathbf{s} \in \mathcal{S}$, such that $p_{n_1}^{(\mathbf{s})} = p_{n_1}^{(\hat{\mathbf{s}})}$ and $p_{n_2}^{(\mathbf{s})} = p_{n_2}^{(\bar{\mathbf{s}})}$, meaning that the actions of users n_1 and n_2 under profile $\mathbf{s}$ result in the minimum cost for them, since $\hat{s}_n$ is a MESE for n_1 and $\bar{s}_n$ is a MESE for n_2. Repeating this process for every user n in the network, i.e., defining the action profile s_n such that $p_n^{(\mathbf{s})} = \min\{p_n^{(\hat{\mathbf{s}})}, p_n^{(\bar{\mathbf{s}})}\}$, we conclude with an action profile $\mathbf{s}$ that yields lower network cost than the MESE points $\hat{\mathbf{s}}$ and $\bar{\mathbf{s}}$. Apart from lower network cost, $\mathbf{s}$ maintains all users' satisfaction, since both $\hat{\mathbf{s}}$ and $\bar{\mathbf{s}}$, from where their actions are selected, are MESE points. Therefore, $\mathbf{s}^* = \mathbf{s}$ is a MESE for game $\mathcal{G}$, which contradicts our initial assumption and proves that the MESE of game $\mathcal{G}$ is unique. ∎

IV. LEARNING SATISFACTION EQUILIBRIA

In this section, we introduce two algorithms to determine the SE and the MESE points. The algorithms are executed in a distributed manner by the different users in the network over several iterations until convergence is met. Each user autonomously selects the action implied by the respective equilibrium type (i.e., SE, MESE) by observing its utility, which is either provided as feedback from the networking environment or calculated separately by each user given the actions selected by its competitors at the previous iteration. In this work, both algorithms assume that the players make simultaneous decisions instantaneously observable by the rest of the players. However, they can be extended to adapt to environments where deviations by users are detected with some time delay by others. For this purpose, locks or timestamps can be used to force the users to play in turns such that only one makes changes to the environment at a time or to specify what actions to take at different time points, respectively.

A. REINFORCEMENT LEARNING-BASED SE ALGORITHM

At the SE, it suffices for each user to satisfy its minimum acceptable tradeoff value without accounting for any optimality, such as the maximization of its utility or the achievement of the minimum possible cost to itself or to the whole network. As a result, no complete knowledge of the actions selected by the rest of the users is required. This gives the opportunity to devise an RL-based algorithm to determine one of the available SE points for the game [9]. The RL algorithm is performed over several iterations, and at each iteration τ, each user selects an available action s_n from its action set $\mathcal{S}_n$ based on some probability distribution over its action space.

To facilitate the analysis, we assign an index $l_n \in \mathcal{L}_n = \{1, \ldots, L_n\}$ to each element in a user's action set $\mathcal{S}_n$, allowing us to arrange them in any desired order. Therefore, we can represent the l_n-th action of user n as s_{n,l_n}. The probability distribution that indicates the preference of each user at iteration τ towards each of its available actions is described by the vector $\boldsymbol{\pi}_n(\tau) = (\pi_{n,1}(\tau), \ldots, \pi_{n,l_n}(\tau), \ldots, \pi_{n,S_n}(\tau))$. For the first algorithm iteration ($\tau = 0$), a uniform distribution is selected, i.e., $\pi_{n,l_n}(0) = \frac{1}{S_n}, \forall s_{n,l_n} \in \mathcal{S}_n$, so as equal probabilities are assigned to all feasible actions of the users. At each of the subsequent iterations, each user's n probability distribution $\boldsymbol{\pi}_n$ is updated by taking into account the satisfaction (or not) of its minimum acceptable tradeoff, the achieved utility, and the current values of the probability distribution vector. Specifically, the update rule is as follows:

$$\pi_{n,l_n}(\tau+1) = \begin{cases} \pi_{n,l_n}(\tau), & \text{if } u_n(\tau) \geq u_n^{thr}, \\ g\big(\pi_{n,l_n}(\tau)\big), & \text{otherwise}, \end{cases} \tag{17}$$

where

$$g\big(\pi_{n,l_n}(\tau)\big) = \pi_{n,l_n}(\tau) + \lambda_\tau r_{n,\tau}\left(\mathbb{1}_{\{s_n(\tau)=s_{n,l_n}\}} - \pi_{n,l_n}(\tau)\right). \tag{18}$$

In this rule, $s_n(\tau)$ is the action selected by user n at iteration τ and $\lambda_\tau = \frac{1}{\tau+1}$ is the learning rate of the algorithm. A high learning rate leads to faster convergence and responsiveness to changes but may also cause instability. Furthermore, $r_{n,\tau}$ is the reward function of the RL process, which is defined as:

$$r_{n,\tau} = \frac{u_n^{max} - u_n(\tau) - u_n^{thr}}{2 \cdot u_n^{max}}, \tag{19}$$

where u_n^{max} is the maximum utility of user n and $u_n(\tau)$ is the utility of user n at iteration τ. Specifically, the term maximum utility u_n^{max} of user n signifies the highest value of utility attained when user n is the sole user associated with its nearest server from a geographical viewpoint, encountering no interference from other users. Meanwhile, the remaining users are associated with the remaining edge servers in a manner that maximizes the overall entropy, i.e., the different sample classes are evenly distributed among the servers. In this way, the rule expressed in (18) intends to allocate greater probabilities to actions that increase the utility value and are more likely to meet the constraint of u_n^{thr}.

The RL algorithm for determining an SE for game $\mathcal{G}$ is presented in Algorithm 1. It is worth noting that the selection of the initial actions can affect its convergence. In particular, the algorithm's convergence depends on the presence of clipping actions for the users that are defined as follows.

Algorithm 1 RL-Based SE Algorithm

Initialize $\tau = 0$;
for each $n \in \mathcal{N}$ in parallel **do**
  for each $s_{n,l_n} \in \mathcal{S}_n$ **do**
    $\pi_{n,l_n}(0) = \frac{1}{S_n}$;
  end for
  Select an action $s_n(0) \sim \boldsymbol{\pi}_n(0)$;
end for
while $\exists n \in \mathcal{N}, u_n(\tau) < u_n^{thr}$ **do**
  for each $n \in \mathcal{N}$ in parallel **do**
    Update distribution $\boldsymbol{\pi}_n(\tau + 1)$ according to (17);
    if $u_n(\tau) \geq u_n^{thr}$ **then**
      $s_n(\tau + 1) = s_n(\tau)$;
    else
      $s_n(\tau + 1) \sim \boldsymbol{\pi}_n(\tau + 1)$;
    end if
  end for
  Update $\tau = \tau + 1$;
end while

Definition 4 (Clipping Action): In the game $\mathcal{G}$, action s_n is a clipping action for user n if and only if:

$$s_n \in f_n(\mathbf{s}_{-n}), \forall \mathbf{s}_{-n} \in \mathcal{S}_{-n}, \tag{20}$$

where $\mathcal{S}_{-n} = \cup_{\forall n' \neq n} \mathcal{S}_{n'}$.

In other words, an action is considered as clipping when the corresponding user remains satisfied upon its selection, regardless of the choices of the other users. The existence of such actions affects the convergence of Algorithm 1 since their selection by a user leads the latter to remain attached to them. To avoid running into infinite loops, Algorithm 1 is terminated after a maximum number of iterations τ_{max} in case convergence has not been achieved up to this point.

Concerning the complexity of Algorithm 1, it depends only on the number of iterations performed until convergence. Specifically, the different steps of the algorithm include algebraic calculations of $\mathcal{O}(1)$ complexity while its execution is performed in an entirely distributed and parallel manner. Furthermore, given that the non-cooperative game among the users is played only once, preceding the FL procedure, the complexity of the FL is not affected by the complexity of the RL-based algorithm and vice versa.

B. BEST RESPONSE DYNAMICS-BASED MESE ALGORITHM

As introduced earlier, the MESE point yields the minimum possible cost for each user separately and the whole network while satisfying all users' minimum acceptable tradeoff value. As a result, each user must address a corresponding minimization problem that necessitates the use of an algorithm based on more deterministic steps. Specifically, the BRD algorithm is used, according to which each user selects the best response action that achieves the minimum cost based on the actions of the rest asynchronously [28].

The devised algorithm comprises two stages that are sequentially executed by the users. To accommodate the description of the first stage, consider that each user's action space $\mathcal{S}_n$ is ordered following the set $\mathcal{L}_n$ defined in the previous section. Then, for each action $s_{n,l_n} \in \mathcal{S}_n$, the user n determines the one action $s_{n,z_n} \in \{s_{n,l_n}, \ldots, s_{n,S_n}\}$ that minimizes its cost, where by $z_n \in \mathcal{L}_n$ we denote the index of this particular action. The index z_n is appended to a corresponding list B_n of user n. Each user initializes its BRD algorithm's execution by selecting as its initial action the first one from the list B_n, i.e., the action that yields the minimum individual cost, $s_{n,B_n[0]}$. The aim of the algorithm is subsequently to select the first action that satisfies the minimum acceptable tradeoff. The latter action is termed as Minimum Satisfying Action (MSA) in the following.

The second stage of the algorithm is performed over several iterations until all users are satisfied. At each iteration τ, each user calculates its MSA denoted as $MSA_n(\tau)$. Owing to the monotonicity of the users' utility function with respect to their experienced cost, i.e., uplink transmission power, it suffices to perform a binary search from the MSA of the previous iteration $MSA_n(\tau - 1)$ to the last action $s_{n,S_n} \in \mathcal{S}_n$. In this way, the user n derives the $MSA_n(\tau)$. Then, the action that incurs the minimum cost to the user from the set $\{s_{n,MSA_n(\tau)}, \ldots, s_{n,S_n}\}$ can be easily determined by the index $B_n[MSA_n(\tau)]$ of the already created list. By utilizing the list B_n, the algorithm prevents each user from performing a binary search over the whole action space $\mathcal{S}_n$, reducing in this way the algorithmic complexity. The overall BRD algorithm is summarized in Algorithm 2. Similarly, with Algorithm 1, if convergence has not been achieved after a maximum number of iterations τ_{max}, the process concludes to prevent potential infinite loops in case the MESE of the game doesn't exist.

Proposition 4: The BRD-based MESE algorithm converges to an action profile $\mathbf{s}^*$ under a finite number of iterations if an SE exists for game $\mathcal{G}$.

Proof: At each iteration of the algorithm, each user examines whether its selected action satisfies its minimum acceptable tradeoff. In case the latter condition is not met, the user is forced to increase its uplink transmission power due to the monotonicity of its utility function with the transmission power level. Thus, a user either sticks to its selected transmission power from the previous iteration or selects a higher one while randomly associating with an edge server that maintains its satisfaction. Given that an SE exists for the game based on Definition 1, the requirements set $f_n()$ will never be empty, and thus, there will always exist the best response to the actions of the rest of the users, securing the algorithm's convergence. ■

Proposition 5: The action profile $\mathbf{s}^*$ concluded by the BRD-based MESE algorithm is a MESE point for game $\mathcal{G}$.

Proof: Let $\mathbf{s}^+$ be an ESE point for game $\mathcal{G}$ and $\mathbf{s}^*$ be the particular point where the BRD algorithm converges, which is also an ESE. To prove that $\mathbf{s}^*$ is a MESE, it should hold true that $p_n^{(\mathbf{s}^*)} \leq p_n^{(\mathbf{s}^+)}, \forall n \in \mathcal{N}$, i.e., the cost incurred to

Algorithm 2 BRD-Based MESE Algorithm

for each $n \in \mathcal{N}$ in parallel **do**
  Initialize $min = c_n(s_{n,S_n})$, $min_{index} = L_n$;
  $B_n[L_n] = L_n$;
  for $l_n = L_{n-1}$ to 1 **do**
    if $c_n(s_{n,l_n}) \leq min$ **then**
      $B_n[l_n] = l_n$;
      $min_{index} = l_n$;
      $min = c_n(s_{n,l_n})$;
    else
      $B_n[l_n] = min_{index}$;
    end if
  end for
end for
Initialize $\tau = 0$, $s_n(0) = s_{n,B_n[0]}, \forall n \in \mathcal{N}$;
while $\exists n, u_n(\tau) < u_n^{thr}$ **do**
  for each $n \in \mathcal{N}$ in parallel **do**
    if $u_n(\tau) < u_n^{thr}$ **then**
      $s_n(\tau + 1) = s_n(\tau)$;
    else
      $MSA_n(\tau) = \text{BinarySearch}(\mathcal{S}_n, MSA_n(\tau - 1))$;
      $s_n(\tau + 1) = s_{n,B_n[MSA_n(\tau)]}$;
    end if
  end for
  Update $\tau = \tau + 1$;
end while

the users at $\mathbf{s}^*$ is lower that at $\mathbf{s}^+$. The proposition will be proved by contradiction. Note that, in this proof, a fixed user-to-edge-server association is considered that is identical for the two ESE points.

Assume that $\exists i \in \mathcal{N}, p_i^{(\mathbf{s}^*)} > p_i^{(\mathbf{s}^+)}$. Given that at each iteration of the algorithm, each user is forced to increase its uplink transmission power, in order for the latter inequality to hold, it means that user i chooses a power level p_i^{exc} that exceeds $p_i^{(\mathbf{s}^+)}$ at some iteration. Denote as $\mathbf{p} = [p_1, \ldots, p_n, \ldots, p_N]$ the powers selected by all users at the exact previous iteration. Then, for user i it holds:

$$p_i \leq p_i^{(\mathbf{s}^+)} < p_i^{exc}, \tag{21}$$

while for the rest of the users, it holds:

$$p_n \leq p_n^{(\mathbf{s}^+)}, \forall n \neq i, n \in \mathcal{N}. \tag{22}$$

Also, the action $p_i^{(\mathbf{s}^+)}$ satisfies the requirements, i.e., $p_i^{(\mathbf{s}^+)} \in f_i(\mathbf{p}_{-i}^{(\mathbf{s}^+)})$, as it constitutes an ESE and thus, an SE. Because of (22) and the fact that the user's i utility increases as the summation of powers of the rest of the users decreases, the power vector $\mathbf{p}_{-i}$ increases the user's i utility and maintains its satisfaction, i.e.,

$$p_i^{(\mathbf{s}^+)} \in f_i(\mathbf{p}_{-i}). \tag{23}$$

Based on the above discussion, we conclude that the power level $p_i^{(\mathbf{s}^+)}$ is the best response, as it satisfies the requirements for the rest of the users' vector $\mathbf{p}_{-i}$ (Eq. (23)) and is lower than p_i^{exc}. This contradicts our assumption that some user i will select a power level p_i^{exc}, such that $p_i^{(\mathbf{s}^*)} > p_i^{\mathbf{s}^+}$ can hold at the point when the algorithm's convergence is reached.

Based on the above:

$$\begin{aligned} p_n^{(\mathbf{s}^*)} \leq p_n^{(\mathbf{s}^+)} &\overset{c(\cdot)\nearrow}{\Longrightarrow} c_n\left(p_n^{(\mathbf{s}^*)}\right) \leq c_n\left(p_n^{(\mathbf{s}^+)}\right), \forall n \in \mathcal{N} \\ &\Longrightarrow \sum_{\forall n \in \mathcal{N}} c_n\left(p_n^{(\mathbf{s}^*)}\right) \leq \sum_{\forall n \in \mathcal{N}} c_n\left(p_n^{(\mathbf{s}^+)}\right), \end{aligned} \tag{24}$$

and the action profile $\mathbf{s}^*$ is a MESE for game $\mathcal{G}$. ■

The complexity of Algorithm 2 is calculated as follows. In the first stage, the action space $\mathcal{S}_n$ of a user n is sorted in ascending order with respect to its uplink transmission power with algorithmic complexity equal to $\mathcal{O}(S_n log(S_n))$. At every iteration of the second stage of the algorithm, the users increase their uplink transmission power or the algorithm converges. In the worst case, this iterative procedure will be repeated for $S = S_1 + \cdots + S_N$ times and user n will be satisfied in $S - S_n$ iterations. At each of the $S - S_n$ iterations, the users perform a binary search to determine their MSA in $\mathcal{O}(S_n log(S_n))$. Given that the algorithm is executed in a fully distributed manner, its overall complexity is $\mathcal{O}((S - S_n) + S_n log(S_n))$. Similar to the RL-based algorithm, the complexity of the BRD-based algorithm does not affect the complexity of the FL and vice versa.

V. EVALUATION & RESULTS

In this section, we evaluate the performance and effectiveness of the proposed framework via modeling and simulation. The simulation topology is initialized as follows. We consider a wireless HFL network arranged within a circular area of 300 m radius. $M = 3$ edge servers are randomly positioned along the perimeter of the circular area, and $N = 10$ users are uniformly distributed within the area, unless otherwise explicitly stated. The total network bandwidth is $W = 10\,\text{MHz}$. The remaining communication parameters are set as: $\rho = -20\,\text{dB}$, $\alpha = 3.5$, $p_{max} = 1\,\text{W}$, and $I_0^m = W_m N_0 = \frac{W}{N} N_0, \forall m \in \mathcal{M}$, where $N_0 = -174$ dBm/Hz [14].

For the implementation of the HFL procedure, we consider $k_1 = 6$ local model updates, $k_2 = 10$ edge model updates, and $K = 100$ overall HFL iterations, while for the training learning rate and the data size of the local model parameters, we assume $\eta = 10^{-3}$, and $Z(\mathbf{w}_n) = 28.1\,\text{Kbits}$, $\forall n \in \mathcal{N}$ [14], [40]. The MNIST dataset [41] is used, and the 60000 total samples are equally divided among the N users following in a non-IID manner by ensuring that no individual user possesses samples of all classes included in the dataset. The local classification models of the users are trained using a commonly used in the literature Convolution Neural Network (CNN) configuration and training, as presented in [42], which consists of two 2-D convolution layers with 5×5 filter size and the Rectified Linear Unit (ReLU) as the activation function. These layers are followed by an additional 2-D Max Pooling layer with pool size 2×2.

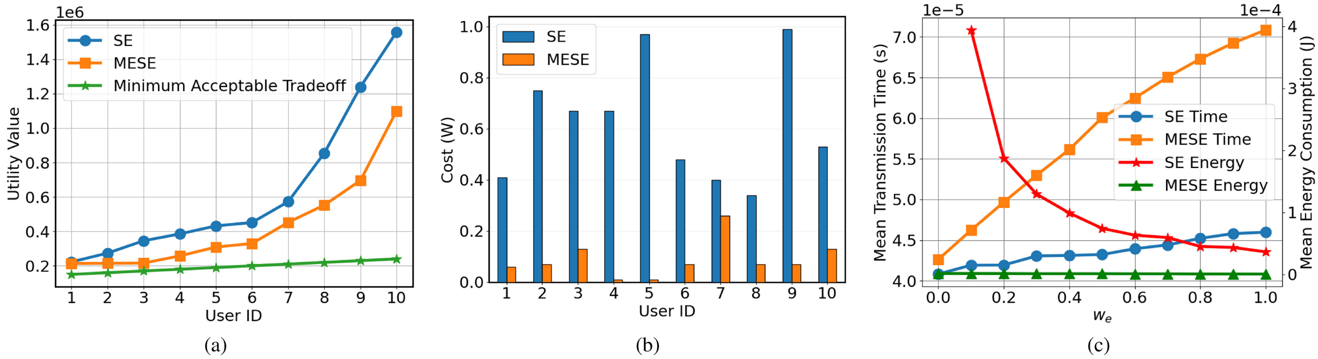

FIGURE 2. **Achieved (a) utility value, (b) experienced cost, and (c) mean consumed energy and time under SE and MESE points.**

Considering the related modeling to the games in satisfaction form, the users' minimum acceptable tradeoff value is defined as $u_n^{thr} = (2 + \frac{n}{10}) \cdot 10^5, \forall n \in \mathcal{N}$, where n is the value of the user index, such that the users' tradeoff requirements become stricter as the user ID increases. The range of values of u_n^{thr} has been experimentally determined to yield reasonable tradeoffs and individual accuracy, time, and energy values, while the aforementioned increasing trend with the user ID has been used for demonstration purposes to facilitate the subsequent numerical results analysis. The convergence limit of the RL-based SE algorithm is set equal to 500 iterations. Also, each user's n action set comprises 100 power levels determined within the feasible range $[0, p_{max}]$ with step 0.01. For statistical purposes, the results presented in the following have been averaged over 100 different initializations corresponding to different data distributions among the users.

A. PERFORMANCE EVALUATION OF DIFFERENT EQUILIBRIA

First, we study the pure operation of the proposed framework and the behavior that the network exhibits based on the type of equilibrium selected for the satisfaction game. Fig. 2(a) illustrates the utility value achieved by each user in the network under the two main types of satisfaction equilibria, i.e., the SE and the MESE, in comparison to the minimum acceptable tradeoff of each user. Note that a high user ID in the horizontal axis indicates a high value for the users' minimum acceptable tradeoff u_n^{thr}, as explained earlier in this section. The numeral results show that independently of the selected equilibrium point, the framework achieves to satisfy all users' minimum acceptable tradeoff value, while a small increase of the latter has as a result a large increase in the users' utility value. Due to the MESE seeking to satisfy the users' tradeoff with the lowest possible cost for them individually and for the network, the achieved utility at this point is lower than the one achieved by the SE. Specifically, given that the cost of each user equals its transmission power and the fact that the utility function u_n is increasing with respect to p_n, the users are forced to lower their satisfaction, maintaining however their tradeoff above its minimum acceptable value.

The results of Fig. 2(a) are further corroborated by Fig. 2(b), where the cost experienced by the users under the two different equilibria is depicted. It is easily deduced that the MESE point concludes with a much lower cost for all users participating in the learning process contrariwise to the SE point that targets exclusively the users' minimum tradeoff value satisfaction. As a result, at the SE point, a remarkable number of the users end up transmitting their local model parameters to the edge with a power level close to its maximum power budget, i.e., p_{max}. Considering the particularly achieved mean consumed energy and time by the users in the network, this is further examined in Fig. 2(c) under different values of the weight factor w_e in the horizontal axis. As expected, a higher value of the weight factor causes a decrease in the energy consumption of the users at the SE point, as higher significance is given to this quantity in the utility function. On the contrary, the users' mean transmission time follows the opposite trend ($w_e + w_t = 1$) at the SE point as the w_e factor increases with a quite slow rate of increase though. Especially regarding the MESE point, the resulting energy consumption is maintained at the lowest possible value regardless of the value of the weight factor w_e as implied by the specific equilibrium point, whereas the mean transmission time of the users increases as w_e gets higher.

Continuing the analysis of the pure operation of the proposed framework, we then evaluate the derived user-to-edge-server association under the two different solution concepts. In more detail, in Fig. 3, the simulated topology is graphically illustrated in terms of the associations derived after the convergence of the RL-based SE (Fig. 3(a)) and BRD-based MESE (Fig. 3(b)) algorithms. The results reveal that the users exhibit a seemingly more random connection pattern to the servers under the SE point, as a means of creating a user data distribution that aligns the network closer to the IID scenario. In contrast, at the MESE point, the users opt for those servers that are in close proximity to them to minimize the incurred transmission cost due to distance and signal attenuation.

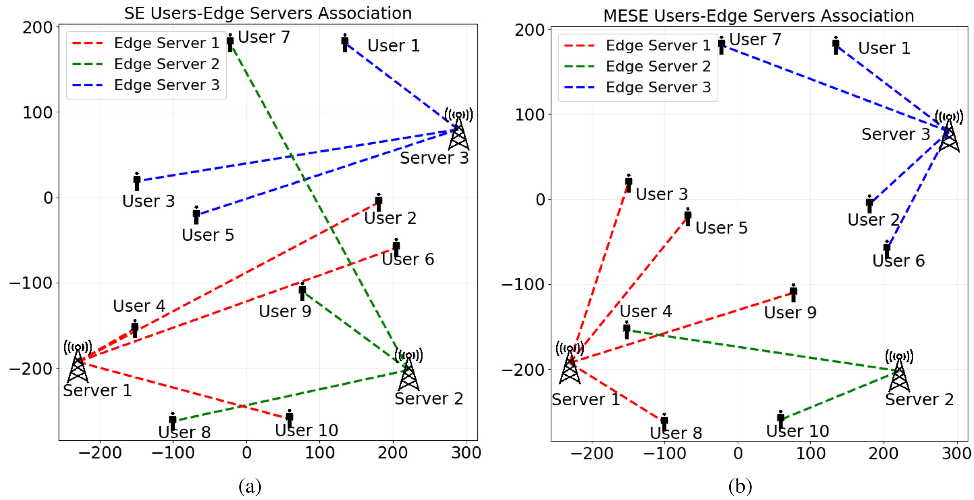

FIGURE 3. Concluded user-to-edge-server association under (a) SE and (b) MESE points.

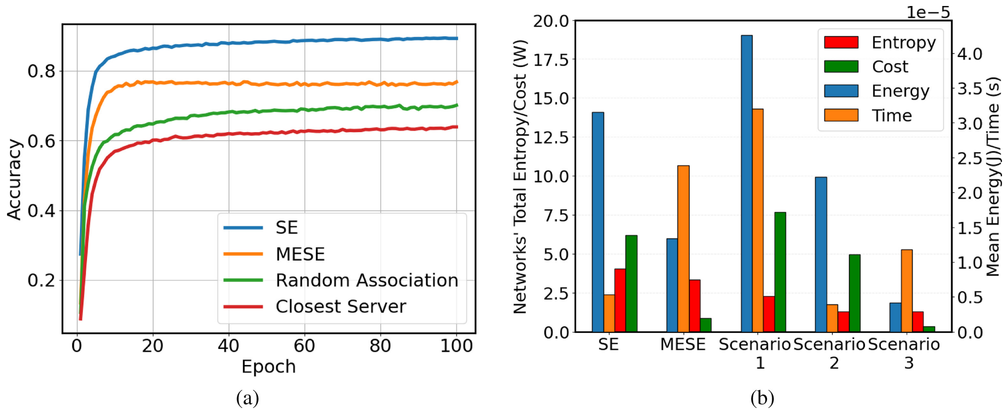

FIGURE 4. Achieved tradeoff in terms of (a) global model's accuracy and (b) network's total entropy, cost, mean consumed energy, and time under SE, MESE, and the benchmark scenarios.

B. PERFORMANCE EVALUATION OF OVERALL PROPOSED FRAMEWORK

Next, we investigate the performance of the proposed satisfaction game-based framework pertaining to the tradeoff achieved between the targeted quantities (networks' entropy, energy consumption, transmission time). For this purpose, we compare the proposed framework against the following three benchmarking scenarios. (i) "Scenario 1": The users randomly select an available edge server to associate with and a power level that lies within the feasible range $[0, p_{max}]$. (ii) "Scenario 2": The users are associated with their closest available edge server and their transmission power is selected as the SE of the game in satisfaction form that is similar to the proposed framework, with the difference that only one variable (i.e., the power) is optimized and considered in the action space $\mathcal{S}_n, \forall n \in \mathcal{N}$. (iii) "Scenario 3": Similar to "Scenario 2" with the difference that the MESE point is selected for the users' uplink transmission power when playing the game in satisfaction form.

Fig. 4(a) depicts the variation of the achieved global model accuracy over the test set as a function of the cloud iterations, i.e., epochs, for both the proposed approach and the benchmarking scenarios. It is important to note that Scenarios 2 and 3 present identical data distribution among the servers due to the user-to-edge-server association based on proximity, which results in identical values regarding the metric of the achieved accuracy too. For this reason, Scenarios 2 and 3 are represented using a single line (red) in Fig. 4(a), labeled as the "Closest Server" scenario. The results reveal that our proposed framework under both types of equilibria (i.e., SE and MESE) attains the highest model accuracy (approximately 80%-90%) among all alternatives, confirming its superior performance in the learning process. The HFL procedure's convergence is achieved after approximately 20 iterations, with the accuracy gradually reaching its final value. Once again it is highlighted that the MESE point sacrifices concluding the best possible accuracy for the sake of its cost minimization objective, yielding a lower accuracy of about 10%. However, still, the proposed framework under the MESE point outperforms all the rest benchmarking scenarios in terms of the achieved model accuracy. Especially for Scenarios 2 and 3, in which the

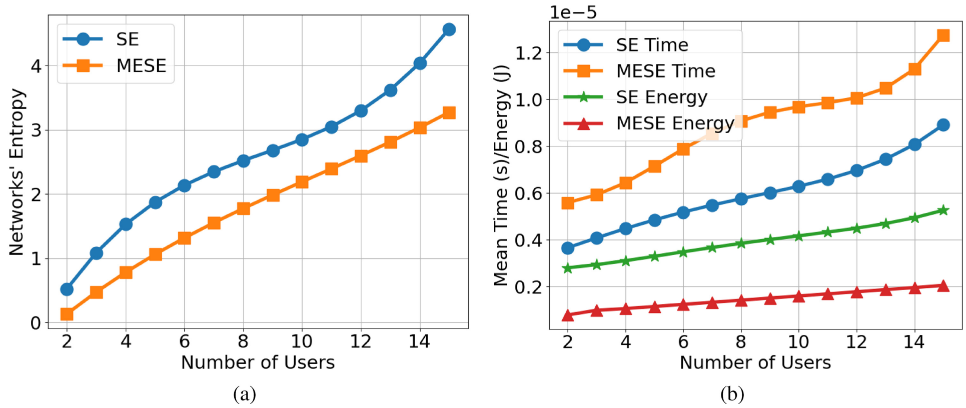

FIGURE 5. Achieved (a) network's entropy and (b) mean consumed energy and time under SE and MESE points for different numbers of users.

users are associated with the nearest server, distant servers are not given the opportunity to collect samples for all the different classes, leading to worse accuracy in the end.

Fig. 4(b) presents a comprehensive overview of the total network's entropy and cost (left axis) and the users' mean energy consumption and transmission time (right axis) under the two types of equilibria of the proposed framework and the three benchmarking scenarios. First, concerning the networks' entropy, we observe that it precisely follows the same trend as the global model's accuracy in Fig. 4(a), with the entropy of Scenarios 2 and 3 having the same value due to the reason already explained. This confirms that the network entropy, capturing the user data distribution across the different edge servers, constitutes a representative metric for the achieved global model's accuracy. Considering the total network cost and energy consumption, the lowest values are achieved under the proposed framework using MESE and Scenario 3 which also considers the MESE point for determining the uplink transmission power allocation. Especially among these two alternatives, Scenario 3 achieves even lower power levels and, thus, energy consumption, as each user selects the closest edge server. However, this comes with the cost of degraded accuracy of the global model, as discussed earlier. Conversely, the proposed framework under SE and Scenario 2 results in a reduction of the required transmission time, which is due to higher transmission power selected by the users, enabling them to achieve higher data rates and thus, lower transmission times. Finally, we can observe that for Scenario 1, where all optimized parameters are randomly selected, the network's performance experiences a substantial degradation and is characterized by notably high energy consumption and transmission time.

C. SCALABILITY EVALUATION

To further examine the impact of the number of network entities in the studied framework in terms of both end users and edge servers, we conduct a scalability evaluation. Specifically, Fig. 5 presents the behavior of the framework as more users are added to the network, ranging from 2 to 15 users. In particular, Fig. 5(a) presents the network's entropy under the proposed framework and the two types of equilibria, i.e., the SE and MESE. Apparently, an increase is shown as more users enter the HFL procedure owing to the inclusion of more (and potentially unique) data samples during model training. Nevertheless, apart from the increase in the network's entropy, an increasing trend is also observed in the mean energy consumption and transmission time of the users as their number gets higher under both SE and MESE (Fig. 5(b)). The increase in the number of users leads to higher congestion in the network and thus, more interference is sensed and caused between them. Consequently, lower data rates are achieved, leading to a longer time required for the transmission of their local model parameters to the edge, along with higher power consumption.

Fig. 6 regards the same network metrics, considering however the case that the number of edge servers increases instead, ranging between 2 and 10 for a fixed number of users equal to $N = 10$. In this simulation case, the trend is entirely the opposite. The presence of a larger number of edge servers implies that the same amount of data samples are shared among more servers, which prevents the network from reaching an IID data distribution case. For small increments in the number of servers, the decrease in entropy is not significant, indicating that the hierarchical structure can provide beneficial support to the overall network, while the rate of decrease of the training's performance is also contingent on the number of users existing in the network (Fig. 6(a)). However, the latter also results in a more uniform distribution of users among the servers, leading to reduced interference and, consequently, lower energy consumption and transmission time (Fig. 6(b)). The comparison between the SE and MESE points results in the same observations as the discussions made so far.

VI. CONCLUSION & FUTURE WORK

In this paper, the joint problem of user-to-edge-server association and uplink transmission power allocation was studied in the radio access part of a wireless HFL network using the NOMA technique. The main objective was to strike an optimal balance for the users participating in the HFL

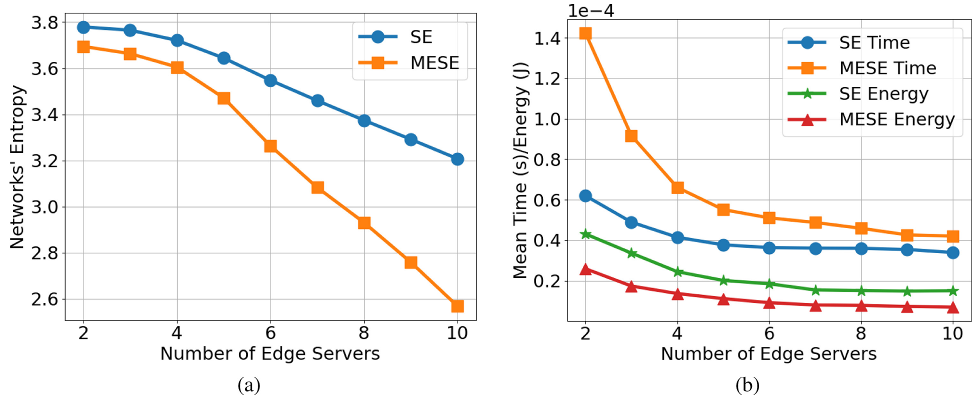

FIGURE 6. Achieved (a) network's entropy and (b) mean consumed energy and time under SE and MESE points for different numbers of edge servers.

procedure, effectively managing the tradeoff between their individual energy consumption, the required transmission time, and the overall performance of the trained global model. In order to address this challenge, a non-cooperative game in satisfaction form was formulated, allowing the different users competing with each other to conclude the respective equilibrium point that satisfies their minimum acceptable accuracy-time-energy tradeoff value. Specifically, different types of equilibria, i.e., SE and MESE, were studied regarding their existence and uniqueness, which take into account different network characteristics, including the users' individual incurred cost and the cost of the whole network. To identify these equilibria, an RL-based and a BRD-based algorithm were devised, respectively. Extensive numerical experiments were conducted to scrutinize the effectiveness of the proposed framework and validate its superiority, concerning the desired energy-time-accuracy tradeoff, comparing also against other benchmarking scenarios.

Part of our current and future work aims to extend the proposed framework to take into account the communications performed at the backhaul network part of the HFL network between the edge servers and the cloud. Furthermore, user mobility and imperfect CSI are crucial factors that introduce stochasticity to the studied HFL network, impacting the proposed solution. Designing Bayesian non-cooperative games to account for these stochasticities is another interesting extension of the current work.

ACKNOWLEDGMENT

The publication of the article in OA mode was financially supported by HEAL-Link.

REFERENCES

[1] Q. Li et al., "A survey on federated learning systems: Vision, hype and reality for data privacy and protection," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 4, pp. 3347–3366, Apr. 2023.

[2] L. U. Khan, W. Saad, Z. Han, E. Hossain, and C. S. Hong, "Federated learning for Internet of Things: Recent advances, taxonomy, and open challenges," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 3, pp. 1759–1799, 3rd Quart., 2021.

[3] Y. Jintao, C. Tan, X. Bowen, S. Yuxuan, Z. Sheng, and N. Zhisheng, "Hierarchical federated learning: Architecture, challenges, and its implementation in vehicular networks," *ZTE Commun.*, vol. 21, no. 1, pp. 38–45, 2023.

[4] N. Mhaisen, A. A. Abdellatif, A. Mohamed, A. Erbad, and M. Guizani, "Optimal user-edge assignment in hierarchical federated learning based on statistical properties and network topology constraints," *IEEE Trans. Netw. Sci. Eng.*, vol. 9, no. 1, pp. 55–66, Jan./Feb. 2022.

[5] S. Liu, G. Yu, X. Chen, and M. Bennis, "Joint user association and resource allocation for wireless hierarchical federated learning with IID and non-IID data," *IEEE Trans. Wireless Commun.*, vol. 21, no. 10, pp. 7852–7866, Oct. 2022.

[6] S. Luo, X. Chen, Q. Wu, Z. Zhou, and S. Yu, "HFEL: Joint edge association and resource allocation for cost-efficient hierarchical federated edge learning," *IEEE Trans. Wireless Commun.*, vol. 19, no. 10, pp. 6535–6548, Oct. 2020.

[7] T. Zhao, F. Li, and L. He, "DRL-based joint resource allocation and device orchestration for hierarchical federated learning in NOMA-enabled industrial IoT," *IEEE Trans. Ind. Informat.*, vol. 19, no. 6, pp. 7468–7479, Jun. 2023.

[8] W. Y. B. Lim, J. S. Ng, Z. Xiong, D. Niyato, C. Miao, and D. I. Kim, "Dynamic edge association and resource allocation in self-Organizing hierarchical federated learning networks," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 12, pp. 3640–3653, Dec. 2021.

[9] S. M. Perlaza, H. Tembine, S. Lasaulce, and M. Debbah, "Satisfaction equilibrium: A general framework for QoS provisioning in self-configuring networks," in *Proc. IEEE Global Telecommun. Conf.*, 2010, pp. 1–5.

[10] P. Promponas, C. Pelekis, E. E. Tsiropoulou, and S. Papavassiliou, "Games in normal and satisfaction form for efficient transmission power allocation under dual 5G wireless multiple access paradigm," *IEEE/ACM Trans. Netw.*, vol. 29, no. 6, pp. 2574–2587, Dec. 2021.

[11] P. Charatsaris, M. Diamanti, and S. Papavassiliou, "On the accuracy-energy tradeoff for hierarchical federated learning via satisfaction equilibrium," in *Proc. 19th Int. Conf. Distrib. Comput. Smart Syst. Internet Things (DCOSS-IoT)*, 2023, pp. 422–428.

[12] D. C. Nguyen, M. Ding, P. N. Pathirana, A. Seneviratne, J. Li, and H. V. Poor, "Federated learning for Internet of Things: A comprehensive survey," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 3, pp. 1622–1658, 3rd Quart., 2021.

[13] Z. Wang, H. Xu, J. Liu, Y. Xu, H. Huang, and Y. Zhao, "Accelerating federated learning with cluster construction and hierarchical aggregation," *IEEE Trans. Mobile Comput.*, vol. 22, no. 7, pp. 3805–3822, Jul. 2023.

[14] L. Liu, J. Zhang, S. Song, and K. B. Letaief, "Client-edge-cloud hierarchical federated learning," in *Proc. IEEE Int. Conf. Commun. (ICC)*, 2020, pp. 1–6.

[15] C. Feng, H. H. Yang, D. Hu, Z. Zhao, T. Q. S. Quek, and G. Min, "Mobility-aware cluster federated learning in hierarchical wireless networks," *IEEE Trans. Wireless Commun.*, vol. 21, no. 10, pp. 8441–8458, Oct. 2022.

[16] Z. Liang, P. Yang, C. Zhang, and X. Lyu, "Secure and efficient hierarchical Decentralized learning for Internet of Vehicles," *IEEE Open J. Commun. Soc.*, vol. 4, pp. 1417–1429, 2023.

[17] Y. Zhou, R. Wang, X. Mo, Z. Li, and T. Tang, "Robust hierarchical federated learning with anomaly detection in cloud-edge-end cooperation networks," *Electronics*, vol. 12, no. 1, p. 112, 2023.
[18] M. H. Mahmoud, A. Albaseer, M. Abdallah, and N. Al-Dhahir, "Federated learning resource optimization and client selection for total energy minimization under outage, latency, and bandwidth constraints with partial or no CSI," *IEEE Open J. Commun. Soc.*, vol. 4, pp. 936–953, 2023.
[19] Y. Zhou, X. Liu, and L. Lei, "Multi-objective optimization for bandwidth-limited federated learning in wireless edge systems," *IEEE Open J. Commun. Soc.*, vol. 4, pp. 954–966, 2023.
[20] X. Ma, H. Sun, and R. Q. Hu, "Scheduling policy and power allocation for federated learning in NOMA based MEC," in *Proc. IEEE Global Commun. Conf.*, 2020, pp. 1–7.
[21] M. S. Al-Abiad, M. Z. Hassan, and M. J. Hossain, "Energy-efficient resource allocation for federated learning in NOMA-enabled and relay-assisted Internet of Things networks," *IEEE Internet Things J.*, vol. 9, no. 24, pp. 24736–24753, Dec. 2022.
[22] H. Chen, S. Huang, D. Zhang, M. Xiao, M. Skoglund, and H. V. Poor, "Federated learning over wireless IoT networks with optimized communication and resources," *IEEE Internet Things J.*, vol. 9, no. 17, pp. 16592–16605, Sep. 2022.
[23] S. M. Perlaza, H. Tembine, S. Lasaulce, and M. Debbah, "Quality-of-service provisioning in decentralized networks: A satisfaction equilibrium approach," *IEEE J. Sel. Topics Signal Process.*, vol. 6, no. 2, pp. 104–116, Apr. 2012.
[24] P. Promponas, E. E. Tsiropoulou, and S. Papavassiliou, "Rethinking power control in wireless networks: The perspective of satisfaction equilibrium," *IEEE Trans. Control Netw. Syst.*, vol. 8, no. 4, pp. 1680–1691, Dec. 2021.
[25] Y. Meng, C. Jiang, L. Xu, Y. Ren, and Z. Han, "User association in heterogeneous networks: A social interaction approach," *IEEE Trans. Veh. Technol.*, vol. 65, no. 12, pp. 9982–9993, Dec. 2016.
[26] A. H. Arani, A. Mehbodniya, M. J. Omidi, and M. F. Flanagan, "Satisfaction based channel allocation scheme for self-organization in heterogeneous networks," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, 2018, pp. 1–6.
[27] A. H. Arani, M. J. Omidi, A. Mehbodniya, and F. Adachi, "Minimizing base stations' ON/OFF switchings in self-organizing heterogeneous networks: A distributed satisfactory framework," *IEEE Access*, vol. 5, pp. 26267–26278, 2017.
[28] P. Promponas, P. A. Apostolopoulos, E. E. Tsiropoulou, and S. Papavassiliou, "Redesigning resource management in wireless networks based on games in satisfaction form," in *Proc. 12th IFIP Wireless Mobile Netw. Conf. (WMNC)*, 2019, pp. 24–31.
[29] H. Li, H. Yu, T. Rao, J. Shang, and F. Shen, "Effects of different loss functions on the classification performance of Hybridsn networks," in *Proc. 3rd Int. Conf. Mach. Learn., Big Data Bus. Intell. (MLBDBI)*, 2021, pp. 500–503.
[30] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Proc. Artif. Intell. Stat.*, 2017, pp. 1273–1282.
[31] M. Diamanti, P. Charatsaris, E. E. Tsiropoulou, and S. Papavassiliou, "Incentive mechanism and resource allocation for edge-fog networks driven by multi-dimensional contract and game theories," *IEEE Open J. Commun. Soc.*, vol. 3, pp. 435–452, 2022.
[32] A. H. Jazi, S. M. Razavizadeh, and T. Svensson, "Integrated access and backhaul (IAB) in cell-free massive MIMO systems," *IEEE Access*, vol. 11, pp. 71658–71667, 2023.
[33] Q. Cheng, Z. Wei, and J. Yuan, "Deep reinforcement learning-based spectrum allocation and power management for IAB networks," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, 2021, pp. 1–6.
[34] S. Zhang, X. Xu, M. Sun, X. Tao, and C. Liu, "Joint spectrum and power allocation in 5G integrated access and backhaul networks at mmWave band," in *Proc. IEEE 31st Annu. Int. Symp. Pers., Indoor Mobile Radio Commun.*, 2020, pp. 1–7.
[35] M. Diamanti, P. Charatsaris, E. E. Tsiropoulou, and S. Papavassiliou, "The prospect of reconfigurable intelligent surfaces in integrated access and backhaul networks," *IEEE Trans. Green Commun. Netw.*, vol. 6, no. 2, pp. 859–872, Jun. 2022.
[36] M. Ma, L. Wu, W. Liu, N. Chen, Z. Shao, and Y. Yang, "Data-aware hierarchical federated learning via task offloading," in *Proc. IEEE Global Commun. Conf.*, 2022, pp. 1–6.
[37] E. E. Tsiropoulou, P. Vamvakas, G. K. Katsinis, and S. Papavassiliou, "Combined power and rate allocation in self-optimized multi-service two-tier femtocell networks," *Comput. Commun.*, vol. 72, pp. 38–48, Dec. 2015.
[38] S. Kakutani, "A generalization of Brouwer's fixed point theorem," *Duke Math. J.*, vol. 8, no. 3, pp. 457–459, 1941.
[39] I. Rival, "A fixed point theorem for finite partially ordered sets," *J. Combinatorial Theory A*, vol. 21, no. 3, pp. 309–318, 1976.
[40] Z. Yang, M. Chen, W. Saad, C. S. Hong, and M. Shikh-Bahaei, "Energy efficient federated learning over wireless communication networks," 2020, *arXiv:1911.02417*.
[41] L. Deng, "The MNIST database of handwritten digit images for machine learning research [best of the Web]," *IEEE Signal Process. Mag.*, vol. 29, no. 6, pp. 141–142, Nov. 2012.
[42] J. Zhang, N. Li, and M. Dedeoglu, "Federated learning over wireless networks: A band-limited coordinated descent approach," in *Proc. IEEE Conf. Comput. Commun.*, 2021, pp. 1–10.

PANAGIOTIS CHARATSARIS received the Diploma degree in electrical and computer engineering from the National Technical University of Athens in 2021, where he is currently pursuing the Ph.D. degree with the School of Electrical and Computer Engineering. His overall research interests lie in the broader area of resource optimization in 5G/6G wireless communications systems.

MARIA DIAMANTI (Member, IEEE) received the Diploma degree in electrical and computer engineering from the Aristotle University of Thessaloniki in 2018, and the Ph.D. degree from the National Technical University of Athens in 2023, where she is a Postdoctoral Researcher with the School of Electrical and Computer Engineering. Her research interests lie in the areas of 5G/6G wireless networks, resource management and optimization, game theory, contract theory, and reinforcement learning.

SYMEON PAPAVASSILIOU (Senior Member, IEEE) is currently a Professor with the School of ECE, National Technical University of Athens. From 1995 to 1999, he was a Senior Technical Staff Member with AT&T Laboratories, NJ, USA. In August 1999, he joined the ECE Department, New Jersey Institute of Technology, USA, where he was an Associate Professor until 2004. He has an established record of publications in his field of expertise, with more than 400 technical journal and conference published papers. His main research interests lie in the area of computer communication networks, with emphasis on the analysis, optimization, and performance evaluation of mobile and distributed systems, wireless networks, and complex systems. He received the Best Paper Award in IEEE INFOCOM 94, the AT&T Division Recognition and Achievement Award in 1997, the U.S. National Science Foundation Career Award in 2003, the Best Paper Award in IEEE WCNC 2012, the Excellence in Research Grant, Greece in 2012, the Best Paper Awards in ADHOCNETS 2015, ICT 2016, IEEE/IFIP WMNC 2019, IEEE Globecom 2022, as well as the 2019 IEEE ComSoc Technical Committee on Communications Systems Integration and Modeling Best Paper Award (for his INFOCOM 2019 paper). He also served on the board of the Greek National Regulatory Authority on Telecommunications and Posts from 2006 to 2009.