Towards Fairness and Green Semantic Communication System: An Anti-discrimination Federated Learning Approach

Guhan Zheng[1], Zhengxin Yu[1]*, Haris Pervaiz[2], Luyao Zhang[3], Haejoon Jung [4], and Syed Ali Hassan [5]
[1]School of Computing and Communications, Lancaster University, UK
[2]School of Computer Science and Electronic Engineering, University of Essex, UK
[3]School of Electronic Information and Communications, Huazhong University of Science and Technology, China
[4]Department of Electronic and Information Convergence Engineering, Kyung Hee University, South Korea
[5]School of Electrical Engineering and Computer Science, National University of Sciences and Technology, Pakistan
*Email of Corresponding Author: z.yu8@lancaster.ac.uk

***Abstract*—Towards addressing emerging energy challenges posed by unfair heterogeneous Semantic Communication (SC) codec updates within future wireless networks, this paper presents a novel Anti-discrimination Federated learning (AdFed) approach. Inspired by the economics of discrimination, unique fairness-associated energy concerns in SC systems are formulated as model discrimination challenges, with the SC-deployed wireless network conceptualized as an anti-discrimination labor market. A novel "affirmative action" strategy, based on training epochs, is proposed and adopted according to historical training unfairness results. To address the reverse discrimination issues in "affirmative action" caused by quota fairness impacting training energy cost, we formulate this problem as a coupled integer non-linear programming problem. Moreover, a new quota trade-off mechanism based on the Rubinstein bargaining game is also designed. Simulation results verify that AdFed outperforms SC training baselines, effectively addressing the unique model discrimination challenges of SC codec model heterogeneity updating. The efficacy of the game theoretical trade-off mechanism is demonstrated in achieving optimal outcomes.**

***Index Terms*—Fairness, energy, semantic communication, anti-discrimination, federated learning, game theory.**

I. INTRODUCTION

In future wireless networks, such as intelligent vehicular networks and aerial-aided networks, Semantic Communication (SC) is recognized as a crucial enabling technology. It allows users to utilize Machine Learning (ML)-based encoders to extract the meaning of information and transmit it to ML-based decoders located at the terrestrial or non-terrestrial base stations deploying edge clouds, enhancing transmission reliability and saving the communication spectrum resource of wireless networks. However, to adapt to new transmission goals and contents, the users' encoders need to be updated in real time in coordination with the decoders at the edges, which introduces a unique challenge. Due to personalization, economic, and privacy concerns, different users may have various encoder models (*i.e.*, encoder models are heterogeneous), while storage-limited edge devices aim to maintain a single decoder model to communicate with multiple users [1].

Federated learning (FL) has been considered a promising approach to tackle the above challenges since it enables users to cooperative training codecs with the edge device under users' data privacy preservation. Several recent research studies have put in the effort to develop FL-based SC systems in general wireless networks [2], [3]. In these proposed systems, users train their SC encoder and decoder models locally for new goals/content and then upload the decoder models to an edge server for model aggregation. To accelerate convergence, various FL-based algorithms tailored for SC codec updating in networks are proposed in [4]–[6]. Nevertheless, the unique heterogeneous codec updating raises a new energy concern that these works fail to consider: **How can we ensure fairness in training within such an SC encoder heterogeneous network and thus avoid wasting training energy**?

Ensuring fairness across groups is crucial in FL studies, as training results are often dominated by data-rich participants. Unfairness leads the vulnerable participants' trained models to be inaccurate, unusable, or denied to be used, thus wasting the training energy of participants.

Two prominent fair FL schemes demonstrate the potential to address this challenge: biased user selection and reweighting federated aggregation. Biased user selection based fair FL [7]–[9] aims to achieve fairness by maximizing the representation of the data distribution through the selection of appropriate users. However, updating SC codecs requires the participation of all necessary users. Without these updates, users can only transmit information via conventional communication paradigms, resulting in degraded network spectrum efficiency. In contrast, the reweighting strategy based fair FL [10]–[12] is to optimize user weighting for federated aggregation based on the amount and type of heterogeneous data. However, most existing fair FL schemes fundamentally focus on data heterogeneity, as it causes the uneven performance of the trained FL model. They lack consideration of model heterogeneity. These data-heterogeneous-based schemes thus cannot work well and encounter new fair challenges in SC encoder model heterogeneous networks.

Unlike users who train homogenous models locally, users train heterogenous models at varying convergence speeds. Existing fair FL schemes make training participation-related decisions based on statistically obtained results during each communication round under incomplete model information. Such information causes unfair training for disadvantaged users due to model differences in the number of neurons and layers, etc. For instance, a type of encoder model may achieve higher accuracy upon training completion, yet converge slowly. Users with another type of faster model convergence may receive higher aggregation weights in the early stage of the training period. This unfair weighting leads to disadvantaged users having inaccurate models and contributing less to federated aggregation, thus perpetuating long-term unfairness. Furthermore, heterogeneous SC codec updating energy challenges if taking measures to preserve fairness, should not be overlooked.

To bridge the gap, we propose a novel Anti-discrimination Federated Learning (AdFed) framework to mitigate the unfairness of SC heterogeneous codec training in networks and save ineffective training energy. Drawing inspiration from the economics of discrimination [13], the SC codec fairness updating problem is formulated as a model discrimination challenge. The wireless network is considered as an anti-discrimination labor market with incomplete information during training. This market includes proposed novel "affirmative action" and corresponding "quotas" to support vulnerable "employees" (users). Further considering energy costs, the energy-related reverse discrimination challenge arising from "quotas" in the context of "affirmative action" is formulated as a coupled integer non-linear programming problem. We thus design a novel Rubinstein bargaining game theoretical approach to trade-off "quotas".

II. SYSTEM MODEL AND ANTI-DISCRIMINATION LABOR MARKET FRAMEWORK

We consider a wireless network where I users employ diverse encoder models while the edge j has only one decoder to receive the information from users. The objective of users is to minimize the training loss of the local encoder model and global decoder model. Hence, we have

$$\{\Theta_i + \Theta_j\} = \operatorname{argmin}\ L_i(\Theta_i + \Theta_j, D_i), \tag{1}$$

where $L_i(\cdot)$ is the loss function and denotes the local objective at user i, D_i is the training data from user i, Θ_i is the encoder model, and Θ_j is the decoder model.

It can be observed that the objective $\{\Theta_i + \Theta_j\}$ of the user i's training is not only related to its encoder model/benefits Θ_i, but also to the decoder model/benefits Θ_j of the edge. Therefore, the relationship between users and the edge can be formulated as a labor market. Users act as employees while the edge functions as the employer. Employees earn benefits via their labor (training) and bring benefits to the employer. By structuring the labor market, we can identify the reasons for unfairness.

Based on the concept of labor market [14], we argue the individual marginal product (local training accuracy) of each employee, $M_i(L_i)$, during training, can be denoted by

$$M_i(L_i) = a_i(D_i) + X_i(\Theta_i), \tag{2}$$

where a_i is the innate ability function related to the number of training data. This is because the relationship between training accuracy and different amounts of data exhibits a logistic-like function for the same training model and investment (epoch) [15]. Differently, X_i represents the human capital function associated with the encoder model. Different training models show no regularity in the training accuracy for the same D_i and epochs. By varying the amount of investment (epoch), different models reveal different non-linear variation trends in accuracy.

Nevertheless, due to privacy concerns, the employer is unable to know about encoder models and data distribution from users, *i.e.*, incomplete information situation, during FL training. The actual individual marginal product believed by the employer can be denoted by [14]

$$M_i^{'}(L_i) = a_i^{'}(D_i) + \varepsilon_i, \tag{3}$$

where $a_i^{'}$ is the employees' innate ability of data (*e.g.*, based on the amount of training data) that employers believe in. It, however, is incomplete information. This is because the distribution of data across users is not independent identically distribution (Non-IID). Further, ε_i is a normal distribution parameter, and the average value is 0. It is associated with employees' model heterogeneity of which the employer is completely unavailable.

Therefore, the employer utilizes this judgment to allocate more benefits (*e.g.*, more aggregation weight) to several employees at the beginning of the work (*i.e.*, FL aggregation). This results in employment statistical discrimination [14]. Furthermore, in the case of conventional fair FL considering only one unfair variable, *i.e.*, Non-IID $a_i^{'}(D_i)$, and only focusing on the unfair during work. The FL information is still incomplete, affecting training decisions/allocation in each epoch. The same long-term model discrimination during training as statistical discrimination [14] in reality thus arose, which results in poor training models and wasted training energy.

Definition 1 (Laissez-faire): Laissez-faire (LF) [16] is an economic concept, meaning people are free to transact without government interference.

Definition 2 (Affirmative action): Affirmative action [17] is the remedial action to equalize the position of different groups in wages, employment, and promotion, based on the historical consequences of inequality.

To combat discrimination from an economic perspective, it is essential to establish a Non-Discriminatory (ND) labor market in place of an LF labor market. The labor market would be less profitable than the ND labor market for both employers and employees if the LF is followed [18]. This is due to some high-cost investments/training being leveled to low-cost lifting. Anti-discrimination benefits both discriminated employees and the employer. To anti-discrimination and establish an ND labor market, “affirmative action” is necessary [17].

III. ADFED: TRAINING PROCESS

To refine the labor market and improve the position of disadvantaged groups, we propose AdFed. The training process of AdFeD is divided into two phases: the “affirmative action” phase and the “distribution according to work” phase. The “affirmative action” aims to enhance the granularity of knowledge sharing between users and the edge, thus mitigating model discrimination caused by incomplete information. In addition, in the ND labor market, users should have equal employment opportunities and receive a fair distribution based on real ability (training data D_i). Hence, after “affirmative action”, AdFed performs “distribution according to work” based on $a_i^{'}(D_i)$ and takes measures to minimize the impacts caused by Non-IID. The training process procedure is outlined in Algorithm 1.

A. *“affirmative action” phase*

The “affirmative action” phase aims to mitigate model discrimination effects, allowing the edge server to focus solely on data heterogeneity (*i.e.*, innate ability). To expedite training, we design a training framework based on split neural network (SplitNN) [15]. Encoder models and the final layer of the decoder model are trained locally by users, while the partial decoder is trained on the edge. Users alternate between training the SC codec model in collaboration with the edge. Compared to FL frameworks that only aggregate identical models, this approach allows heterogeneous encoder models to learn global information and therefore accelerates convergence, as discussed in one of our previous papers [15]. The progress of “affirmative action” is detailed as follows:

1) : Each user global trains x epoch while the edge collects historical loss values and calculates the corresponding benefit based on these loss values.

2) : Before $x + 1$ epoch, the edge calculates the Gini coefficient between each user’s average benefit and the user with the highest average benefit. According to sociological studies, a Gini index greater than 0.4 indicates that the value of benefits tends to be unfair, necessitating “affirmative action”. Low-benefit users should receive more training resources.

3) : Our “affirmative action” is implemented by adaptively adjusting the affirmative quota, *i.e.*, local training epoch. Low-benefit users train with their own encoder model along with the global decoder for a specified “quota”. At the end of “affirmative action”, *i.e.*, after additional training, the encoder model is preserved while the decoder model is discarded to prevent it from memorizing too much of the low-benefit user’s data distribution. Details on quota determination will be elaborated in the next section.

4) : After “affirmative action”, the new encoder model of users performs $x + 1$ epoch global training with the decoder model from x epoch on the edge.

5) : Following the $x+1$ epoch, the users and edge proceed with the next x epochs.

Algorithm 1 AdFed

```
Input: dataset ${D_1, D_2, ..., D_i}$
Output: dataset ${D_1^j, D_2^j, ..., D_i^j}$ from $\Theta_j$'s output
"affirmative action" phase:
Batch size: $K$
for each training epoch $\varphi$
  for each user $i$ involved in training $i = 1, 2, ...I$
    User and edge collaboration training
    Loss $L_i^{\varphi} \longleftarrow \frac{1}{K}\sum_{k=1}^{K}(D_{i,k} - D_{i,k}^{j})$
    Average loss $L_{i,ave}^{\varphi} = \frac{\sum_{\omega=\varphi-x}^{\varphi/x} L_i^{\omega}}{\varphi/x}$
    Update encoder as $\Theta_i^{\varphi}$ and decoder as $\Theta_j^{\varphi}$
  end for
  if $\varphi \% x = 0$
  for each user $i$ involved in training $i = 1, 2, ...I$
    Gini index = $\frac{L_{i,ave}^{\varphi} - L_{i,max}^{\varphi}}{L_{i,ave}^{\varphi} + L_{i,max}^{\varphi}}$
    if Gini index>0.4
      Gaming the quota $n$
      for each training epoch $\ell$ in $n$
        Loss $L_i^{\ell} \longleftarrow \frac{1}{K}\sum_{k=1}^{K}(D_{i,k} - D_{i,k}^{j})$
      end for
      Update the encoder as $\Theta_i^{\varphi+n}$ and keep the decoder as $\Theta_j^{\varphi}$
  end for
end for
Distribution according to work phase:
Batch size: $K$
All users predicted loss value decreasing gradient less than the threshold based on Eq. (6)
for each training epoch $\varphi$
  for each user $i$ involved in training $i = 1, 2, ...I$
    User local training
    Loss $L_i^{\varphi} \longleftarrow \frac{1}{K}\sum_{k=1}^{K}(D_{i,k} - D_{i,k}^{j})$
    Update encoder as $\Theta_i^{\varphi}$ and decoder as $\Theta_{i,j}^{\varphi}$
    Uplaod $\Theta_{i,j}^{\varphi}$ to the edge
  end for
  edge perform fair federated aggregation
  edge delivery the aggregated global decoder $\Theta_j^{\varphi}$ to users
end for
```

B. *Distribution according to work phase*

However, the SplitNN-based approach is inherently unfair for various data distribution users because the decoder tends to favor the last user trained [19]. Therefore, when all users’ training tends to converge, the distribution according to the work phase begins. In this phase, users train the full codec model locally (*i.e.*, both encoder and decoder) and upload only the decoder model for federated aggregation. The goal of federated aggregation is to ensure data training fairness and reduce the influence of so-called vulnerability due to training order in SplitNN. The effects of model discrimination can be overlooked due to the presence of “affirmative action”, allowing the training to focus on data heterogeneity, *i.e.*, non-IID. Since data heterogeneity is not the focus of this paper, introduced fair FL approaches for non-IID, *e.g.*, [10]–[12] can be adopted.

IV. ADFED: QUOTA TRADE-OFF

Excessive quotas would result in unfair reverse discrimination (high energy cost). A small quota would fail to act against discrimination. Therefore, in this section, based on the ND labor market, to strive for non-discrimination and save energy, we analyze the economic utilities for employees and employer separately, followed by the presentation of our solution.

A. Economic utility of employees (users)

Without loss of generality and in order to save energy, the same as [5], we consider the training energy consumption as the monetary cost. The economic utility achievable via "affirmative action" of the user i thus should be anti-discrimination income minus the cost and old income under LF [18]. It can be denoted by

$$u_i(n) = B_i^{ND}(n) - B_i^{LF} - E_i^c(n) - E_i^t(n), \tag{4}$$

where n is the number of "affirmative actions", *i.e.*, epochs. Further, B_i^{LF} is the benefit of user i in the LF labor market. Here, considering the loss value is the benefit during training and generally decreases with training epochs. Inspired by the accuracy metric, *e.g.*, eak signal-to-noise ratio (PSNR), the benefit/accuracy during the work B_i of user i's loss can be formulated via the logarithmic function. We have

$$B_i(L_i)^{LF} \triangleq \eta \log_{10}(\frac{1}{L_i}), \tag{5}$$

where η is the monetary parameter. Further, E_i^c and E_i^t are the computing energy cost and the transmission cost for "affirmative action", separately.

In addition, B_i^{ND} is the benefit of user i after "affirmative action", *i.e.*, work in the ND labor market. Here, the benefit $B_i^{ND} = \eta \log_{10}(\frac{1}{L_i(n)})$ after "affirmative action" is a function related to epochs. Therefore, we need to predict and formulate the real-time $L_i(n)$ function for B_i^{ND}. As $L_i(n)$ and n can be considered as a binary relation, least squares fitting is an efficient method. The target of least squares fitting can be expressed by

$$\min \sum_{m=1}^{M+1} (L_i^{n_m} - L_i(n_m)), \tag{6}$$

where M is the number of epochs trained in this historical training epoch and $+1$ is the final expected loss value, $L_i^{n_m}$ is the true output loss value during training and $L_i(n_m)$ is the fitting function of L_i via the least squares method.

The energy cost E_i^c and E_i^t can be expressed as:

$$E_i^c + E_i^t = n(D_i \kappa a_i f_i^2 + \frac{D_i p_i}{r_i}), \tag{7}$$

where κ is the CPU-cycle related parameter, a_i is the encoder model size factor, f_i is the CPU-cycle, p_i is the transmission power and r_i is the transmission rate.

Therefore, the optimal objective quota of users can be expressed by $\max_n u(n)$.

B. Economic utility of the employer (edge)

Similar to users, for the edge, the economic utility is also denoted by income achievable via "affirmative action" minus energy cost. We have the utility as:

$$U(n) = B_i^{ND}(n) - B_i^{LF} - E_j^c - E_j^t - S_i, \tag{8}$$

where E_j^c and E_j^t are the computing energy cost and the transmission cost of edge j for "affirmative action", separately. We have

$$E_j^c + E_j^t = n(D_i \kappa a_j f_j^2 + \frac{D_i p_j}{r_j}), \tag{9}$$

where κ is the CPU-cycle related parameter, a_j is encoder model size, f_i is the CPU-cycle, p_i is the transmission power and r_i is the transmission rate.

In addition, different from users, to avoid reverse discrimination, we incorporate the change in satisfaction of other users as the cost into the edge utility, as in some game-theoretic studies, *e.g.*, [20]. It is the result of the increasing "affirmative action" and is denoted by S_i in Eq. (8). The change in satisfaction is mainly related to the training delay, which can be expressed as:

$$S_i = \begin{cases} \iota \ln(1 + \vartheta - T_i), & \text{if } \vartheta \geq T_i \\ 0, & \text{if } \vartheta < T_i \end{cases} \tag{10}$$

where ι is the monetary parameter and ϑ is the maximum tolerable delay. Further, T_i is the "affirmative action" delay. It can be denoted by

$$T_i = n(\frac{D_i a_i}{f_i} + \frac{D_i}{r_i}) + n(\frac{D_i a_j}{f_j} + \frac{D_i}{r_j}). \tag{11}$$

Therefore, the objective of the "affirmative action" performed by the edge can be denoted by $\max_n U(n)$.

C. Rubinstein bargaining game

We can observe that for the same action n, the edge and users should have different optimal quotas due to different objectives. These two optimization problems constitute coupled nonlinear integer programming problems. Based on the economic labor market perspective, we formulate the edge and the user to play an economic Rubinstein bargaining game to approximate the Nash Equilibrium (NE) point.

The user and edge play both sides of the game and want to maximize their utility via bargaining their offer. The user can provide their offer, which can be denoted by

$$\gamma_i^t u_i^{max}(n) = \gamma_i^t \max_n u_i(n), \tag{12}$$

where t is the times of offers and γ_i^t dissipation factor of user i. It shows la onger bargaining duration due to an increase in the times of offers, which leads to a decrease in utility. Similarly, the edge offer can be expressed as:

$$\gamma_j^t U^{max}(n, \gamma_j^t) = \gamma_j^t \max_n U(n, \gamma_j^t), \tag{13}$$

where $U(n, \gamma_j^t)$ is affected by the times of bargains due to the existence of satisfaction consideration.

With continued offers in time t, both parties can reach NE. However, it requires a continued n to achieve game-perfect NE and increases the complexity of computing the optimal u^* and U^*. Since epochs (quotas) are integer and discontinuous, we propose a one-step bargain to approximate the game-perfect NE. We found that in our scenario, the optimal decision must be between the first offers ($u^* \leq \gamma_i^1 u^{max}(n) \& U^* \leq \gamma_j^1 U^{max}(n)$) and the game aims to maximize social welfare, *i.e.*, $\max U + u$ Therefore, we utilize the two-dimensional grid search for the approximate NE that maximizes $U + u$ from a limited set of integers, in the case of only one communication/bargain.

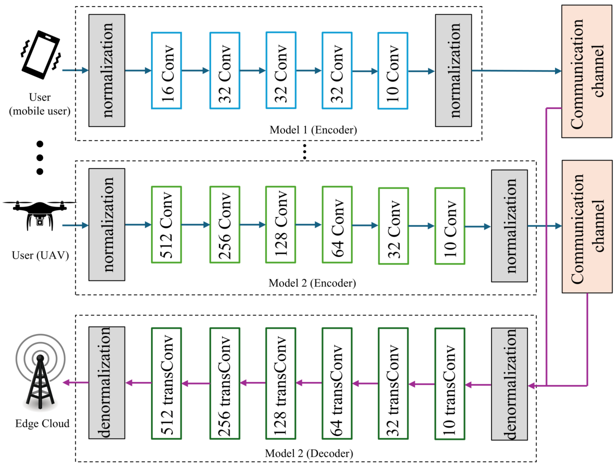


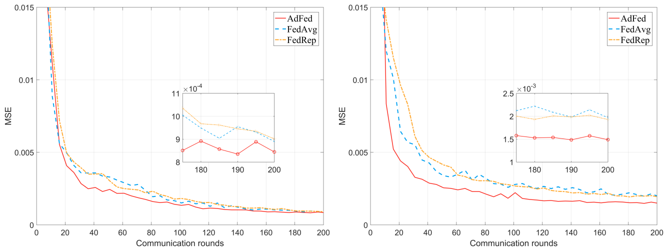
Fig. 1. Encoder and decoder neural network architectures used in the simulations.

V. SIMULATIONS

In the simulation, we choose two types of different-sized SC codec frameworks as benchmarks for image transmission from existing SC studies: a small model [21], referred to as Model 1, and a relatively bigger model [4], referred to as Model 2. The network environment is simulated with five users transmitting images to an edge. Due to page limitations, we choose the two most popular image datasets, CIFAR 10 and CIFAR 100, as the dataset for simulation. To visualize the advantages of our frameworks, we configure four users (e.g., mobile users) to use the Model 1 encoder and one user (e.g., uncrewed aerial vehicle (UAV)) to use the Model 2 encoder, with the edge deploying a Model 2 decoder (Fig. 1). The users' data are non-IID, the Model 2 user has more data information than others, and "Affirmative action" is implemented every 5 epochs. Furthermore, we consider the feasibility of the SC scenario, compare our AdFed with only-data-aware fair FL based on a half-model aggregation framework, *i.e.,* FedRep [22], and with FedAvg [23] for the benchmark disregarding the model heterogeneity. The accuracy metric is the mean-square error (MSE) of the test set during training [15].

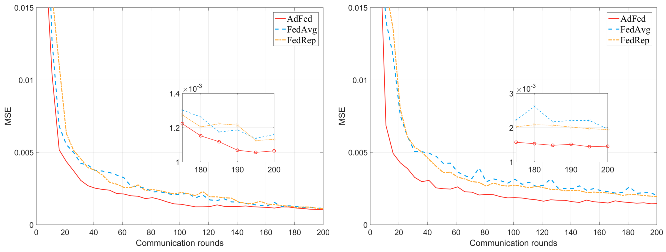
TABLE I
GINI INDEX

	FedAvg	**FedRep**	**AdFed**
CIFAR 10	0.3346	0.4143	0.2255
CIFAR 100	0.2915	0.3052	0.1887



(a) Model 1 (b) Model 2

Fig. 2. Convergence speed and accuracy of various learning algorithms with CIFAR 10 datasets.



(a) Model 1 (b) Model 2

Fig. 3. Convergence speed and accuracy of various learning algorithms with CIFAR 100 datasets.

Fig. 2 illustrates the effectiveness of our scheme under the CIFAR10 dataset. It can be observed that from the comparison of Fig. 2 (a) and Fig. 2 (b), Model 2 is less accurate than Model 1 based on FedAvg. Hence, there exists a certain issue of the unfairness of the heterogeneous model. We can then observe from Fig. 1 (a) and Fig. 1 (b) that with AdFed, both Model 1 users and Model 2 users achieve faster convergence and higher accuracy compared to the baselines. The advantage for the Model 2 user, who was initially discriminated against, becomes more pronounced. It narrowed the unfair gap between Model 1 and Model 2. This improvement is attributed to the energy-aware anti-discrimination measures in AdFed, which allow the additional information carried by the Model 2 user to be effectively learned, thereby enhancing the global model during training. We then change the dataset to CIFAR 100 and show the results in Fig. 3. We can see that Fig. 3 shows the same trend and demonstrates the effectiveness of our AdFed.

Table I depicts our improvements in fairness. The Gini index is used to measure fairness. It can be observed that our AdFed consistently achieves the lowest Gini index. This proves that our AdFed is effective in mitigating unfairness.

TABLE II
PROPORTION OF ENERGY WASTED

Trained encoder unavailable	n=1	n=3	n=5
100%	10%	30%	50%

This is because during the "affirmative action" phase, in case there exists great unfairness, our scheme performs "affirmative action" to mitigate the inequity. Therefore, the probability of the semantic codec being unavailable is reduced, thus saving training energy resources.

In Table II, we show the network energy saved through our proposed AdFed. We substitute the number of neurons being trained for the energy consumed. If the accuracy of the trained model does not meet the user's criteria, then we regard it as 100% of the training resources being wasted. Because of the stochastic nature of training, we use the average increase in the number of "affirmative actions", n, to represent the additional energy savings. It can be observed that even in the case of 5 more training sessions per affirmative action, the extra energy expended by the affirmative action is only half of the completed training. However, our scheme increases the fairness as well as the accuracy of the training and thus reduces the discard rate of the trained model.

VI. CONCLUSION

In this paper, we explored the fairness-related energy challenges associated with SC codec updating in encoder heterogeneous networks and proposed a pioneering framework named AdFed. AdFed employs the concepts of the economics of discrimination to structure the SC encoder heterogeneous network as an AD labor market, mitigating the SC unfairness challenge, *i.e.*, the model discrimination. "affirmative action" for fair SC codec updating, based on training epochs, was further proposed. Furthermore, energy-aware reverse discrimination and SC technical challenges within AdFed were scrutinized. To tackle these challenges, a new "quota" trade-off mechanism based on the Rubinstein bargaining game was presented. The simulation results demonstrate the superiority of our proposed AdFed framework and associated approaches.

VII. ACKNOWLEDGMENTS

This work was supported by the MSIT, Korea, in part under the National Research Foundation of Korea grants (RS-2023-00303757), in part under the ITRC support programs (IITP-2025-RS-2021-II212046), and in part under the Convergence security core talent training business support program (IITP-2023-RS-2023-00266615) supervised by the IITP.

REFERENCES

[1] G. Zheng, Q. Ni, A. Kaushik, and L. Yang, "Semantic communication meets heterogeneous network: Emerging trends, opportunities, and challenges," *arXiv preprint arXiv:2502.08999*, 2025.

[2] Z. Qin, G. Y. Li, and H. Ye, "Federated learning and wireless communications," *IEEE Wireless Communications*, vol. 28, no. 5, pp. 134–140, 2021.

[3] H. Tong, Z. Yang, S. Wang, Y. Hu, O. Semiari, W. Saad, and C. Yin, "Federated learning for audio semantic communication," *Frontiers in communications and networks*, vol. 2, p. 734402, 2021.

[4] G. Zheng, Q. Ni, K. Navaie, and H. Pervaiz, "Semantic communication in satellite-borne edge cloud network for computation offloading," *IEEE Journal on Selected Areas in Communications*, vol. 42, no. 5, pp. 1145–1158, 2024.

[5] G. Zheng, Q. Ni, K. Navaie, H. Pervaiz, A. Kaushik, and C. Zarakovitis, "Energy-efficient semantic communication for aerial-aided edge networks," *IEEE Transactions on Green Communications and Networking*, pp. 1–1, 2024.

[6] G. Zheng, Q. Ni, K. Navaie, and C. Zarakovitis, "Socially-inspired semantic communication codec updating for ntn-enabled intelligent transportation systems," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–13, 2025.

[7] T. Nishio and R. Yonetani, "Client selection for federated learning with heterogeneous resources in mobile edge," in *ICC 2019-2019 IEEE international conference on communications (ICC)*. IEEE, 2019, pp. 1–7.

[8] M. Yang, X. Wang, H. Zhu, H. Wang, and H. Qian, "Federated learning with class imbalance reduction," in *2021 29th European Signal Processing Conference (EUSIPCO)*. IEEE, 2021, pp. 2174–2178.

[9] C. Shi and C. Shen, "Federated multi-armed bandits," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 35, no. 11, 2021, pp. 9603–9611.

[10] A. Abay, Y. Zhou, N. Baracaldo, S. Rajamoni, E. Chuba, and H. Ludwig, "Mitigating bias in federated learning," *arXiv preprint arXiv:2012.02447*, 2020.

[11] Y. Zeng, H. Chen, and K. Lee, "Improving fairness via federated learning," *arXiv preprint arXiv:2110.15545*, 2021.

[12] Y. H. Ezzeldin, S. Yan, C. He, E. Ferrara, and A. S. Avestimehr, "Fairfed: Enabling group fairness in federated learning," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 37, no. 6, 2023, pp. 7494–7502.

[13] G. S. Becker, *The economics of discrimination*. University of Chicago press, 2010.

[14] S. J. Lundberg and R. Startz, "Private discrimination and social intervention in competitive labor market," *The American economic review*, vol. 73, no. 3, pp. 340–347, 1983.

[15] G. Zheng, Q. Ni, K. Navaie, H. Pervaiz, G. Min, A. Kaushik, and C. Zarakovitis, "Mobility-aware split-federated with transfer learning for vehicular semantic communication networks," *IEEE Internet of Things Journal*, vol. 11, no. 10, pp. 17 237–17 248, 2024.

[16] J. M. Keynes, "The end of laissez-faire," in *Essays in persuasion*. Springer, 1926, pp. 272–294.

[17] J. Rubenfeld, "Affirmative action," *Yale LJ*, vol. 107, p. 427, 1997.

[18] S. J. Lundberg, "The enforcement of equal opportunity laws under imperfect information: affirmative action and alternatives," *The Quarterly Journal of Economics*, vol. 106, no. 1, pp. 309–326, 1991.

[19] H. Madaan, M. G. V. Kulkarni, and A. Pant, "Vulnerability due to training order in split learning," in *ICT Systems and Sustainability: Proceedings of ICT4SD 2021, Volume 1*. Springer, 2022, pp. 103–112.

[20] Y. Wang, W. Chen, T. H. Luan, Z. Su, Q. Xu, R. Li, and N. Chen, "Task offloading for post-disaster rescue in unmanned aerial vehicles networks," *IEEE/ACM Transactions On Networking*, vol. 30, no. 4, pp. 1525–1539, 2022.

[21] E. Bourtsoulatze, D. B. Kurka, and D. Gündüz, "Deep joint source-channel coding for wireless image transmission," *IEEE Transactions on Cognitive Communications and Networking*, vol. 5, no. 3, pp. 567–579, 2019.

[22] L. Collins, H. Hassani, A. Mokhtari, and S. Shakkottai, "Exploiting shared representations for personalized federated learning," in *International conference on machine learning*. PMLR, 2021, pp. 2089–2099.

[23] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Artificial intelligence and statistics*. PMLR, 2017, pp. 1273–1282.