2024 6th International Conference on Communications, Signal Processing, and their Applications (ICCSPA) | 979-8-3503-8481-9/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICCSPA61559.2024.10794379

Trustworthy Aggregation for Aerial Federated Learning in Heterogeneous Client Environments

Mohamed Ads, Student Member, IEEE, Hesham ElSawy, Senior Member, IEEE,
Hazem M. Abbas, Senior Member, IEEE, Hossam S. Hassanein, Fellow, IEEE
School of Computing, Queen's University, ON, Canada
m.ads@queensu.ca, hesham.elsawy@queensu.ca, hazem.abbas@queensu.ca, hossam@cs.queensu.ca

***Abstract*—The integration of unmanned aerial vehicles (UAVs) within Federated Learning (FL) marks a crucial advancement, introducing a dynamic and flexible dimension to decentralized machine learning paradigms. However, the location-dependent performance, characterized by variations in transmission rates and susceptibility to errors, presents challenges for FL's convergence speed and accuracy. This paper proposes a trustworthy aggregation approach implemented on non-terrestrial networks, addressing client heterogeneity in aspects of trustworthiness, available datasets, and transmission rates. The proposed approach initially prioritizes clients with high data rates is implemented to expedite convergence speed, gradually expanding to incorporate a more extensive client base. This accommodation of clients is pivotal for training the model on a larger decentralized dataset that is a crucial consideration in scenarios with non-independent and identically distributed datasets. However, the potential presence of imperfect local models stemming from small datasets, undetected attacks, or less powerful computational devices may result in a gradual degradation of training performance over time. In such instances, the system exclusively focuses on training with a reliable set of clients. The proposed algorithm is subjected to benchmarking against two scenarios: an aggressive approach, which involves accommodating all clients, and a conservative approach, which only includes authenticated clients. Notably, the proposed algorithm excels in both cases, especially in environments with lower levels of trust.**

I. Introduction

The advent of sixth-generation technology, referred to as 6G, has become a subject of significant interest and scholarly inquiry. Positioned as the successor to 5G, 6G stands poised to revolutionize connectivity by promising unprecedented data speeds, ultra-low latency, and groundbreaking advancements in network capabilities. Concurrently, unmanned aerial vehicles (UAVs) have emerged as versatile tools with applications spanning various disciplines. The convergence of 6G and UAV technologies marks a pivotal juncture, presenting synergies that hold the potential to redefine paradigms within the domain of wireless communication.

The integration of Federated Learning (FL) [1] into the 6G framework adds a layer of sophistication to this evolving landscape. This relationship between FL and 6G technologies takes advantage of UAVs, elevating FL capabilities by accommodating an expanded network of clients. The synergy between FL and UAVs not only extends the reach to a larger clientele but also ensures the transmission of models with superior signal quality [2], marking a significant advancement in the convergence of 6G and FL technologies. As these innovative collaborations unfold, they hold the promise of revolutionizing the landscape of data transmission and machine learning in the forthcoming era of advanced wireless communication. One of the most used algorithms in aggregating the local weights is Federated Averaging (FedAvg) [3]. This algorithm performs a weighted sum based on the size of the dataset for each client.

The success of the aforementioned aggregation algorithm is contingent upon the Signal-to-interference-plus-Noise Ratio (SINR) of the clients. In instances where the (SINR) falls below optimal levels, the performance of the FedAvg algorithm may be adversely affected because the client will not participate in the training process and the model will become more biased toward clients with high signal quality. This nuanced consideration of environmental factors underscores the need for a comprehensive understanding of the real-world conditions in which these algorithms operate. In addition, the algorithm assumes accurate and true models are being uploaded by the clients. However, in real-world environments, assuming every participating client operates with good intentions and consistently delivers accurate model updates is not practical.

The authors in [4]–[6] developed a technique to counter the biasing effect inherent in FedAvg due to participation failure. Moreover, in [5], the author analyzed the biasing of FedAvg in non-terrestrial networks but did not take into consideration the accuracy and security of the client's participation. On the other hand, authors in [7] explored the biasing effect and security aspects in only terrestrial environments.

The authors in [8], [9] proposed secure FL that introduces a ranking of the devices based on their historical participation. However, they did not take into consideration the critical wireless communication challenges, including global model bias and a sluggish convergence rate. To the best of the authors' knowledge, this paper stands as the pioneering effort in considering the combined influence of trustworthiness and wireless network characteristics on FL. An FL model with security considerations is presented in [8], [9], incorporating a trustworthiness ranking based on devices' past participation. Nevertheless, they neglect significant challenges in wireless communication, such as global model bias and a slow convergence rate.

To the best of our knowledge, our work is the first paper to comprehensively examine the joint effects of trustworthiness and wireless network characteristics on FL within non-

terrestrial settings as follows:

1. Developing a weighting factor that counters the impact of non-terrestrial impairments (i.e. fading, interference)

2. Exploring the influence of clients with drifted models on FL in environments with limited data sources.

The results underscore the importance of considering client trustworthiness, as overlooking this factor results in significant model degradation. Also, solely depending on authenticated clients is inadequate due to limited data availability. The proposed hybrid model achieves a distinctive equilibrium between precision and convergence speed by engaging uncertainty clients, with moderate trustworthiness scores, in the initial training rounds and discarding them once they negatively affect the model.

II. System Model and Problem Formulation

A cellular network is designed with a single tier, utilizing unmanned aerial vehicles (UAVs) acting as aggregators deployed based on a Poisson point process (PPP). Around each UAV, a set of clients C is uniformly distributed such that each client connects to the closest UAV geographically, as illustrated in the figure. For simplicity, we focus on a representative UAV placed at an arbitrary origin and its associated clients represented by the set.

A. Federated Learning Model

Within the Client set C, each client is assigned an index i and possesses an individual dataset denoted as D_i for updating its local parameters represented by $\boldsymbol{w}_{i,t}$. The adjustment of local weights follows the iterative formula:

$$\boldsymbol{w}_{i,t}^{(x+1)} = \boldsymbol{w}_{i,t}^{(x)} - \eta g_i(\boldsymbol{w}_{i,t}^{(x)}) \tag{1}$$

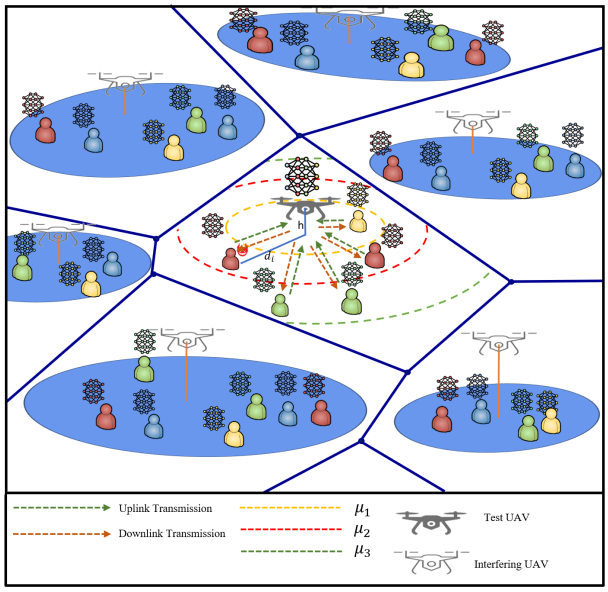

Fig. 1. A dynamic SINR based aggregation implemented on a UAV with 3 levels and five clients

In this context, $\boldsymbol{w}_{i,t}^{(x+1)}$ signifies updated weights post stochastic gradient descent with learning rate η, step size ψ, and momentum β at the global round t, within the epoch range of 1 to the total local epochs E. Subsequently, all clients will transmit $\boldsymbol{w}_{i,t}$ which are the weights of the model after performing the last epoch $\boldsymbol{w}_{i,t}^{(E)}$ to their associated UAV to aggregate them. The global weight $\boldsymbol{g}_t$ is then computed through the FedAvg algorithm and redistributed to all participating clients. A schematic representation of the system model is illustrated in Fig. 1.

Throughout the iterative updating process, the server consolidates the distributed optimization task using the following expression:

$$\min_{\boldsymbol{w}} f(\mathbf{g}) = \frac{1}{U}\sum_{n=1}^{U} f_i(\boldsymbol{w}),$$

here, $f_i(.)$ refers to the loss function associated with client i. Additionally, it is assumed that there is a validation set that is accessible by the UAV such that it evaluates the accuracy of the aggregated model [10] [11].

B. Communication Model

The network is assumed to implement a universal frequency reuse strategy, utilizing orthogonal channels in which individual clients are allocated specific resource blocks (RBs). This approach facilitates simultaneous data transmission without casing interference inside the cell by ensuring that the number of RBs is either equal to or greater than the number of devices within each cell. However, a challenge emerges as interference may originate from devices coexisting within the same RB but located on other cells.

Considering the wireless communication environment, we assume that all clients transmit with the same power P. Moreover, to model the impact of obstruction in the surroundings, we employ the estimated probability of Line of Sight (LOS) as detailed in [12]. This probability characterizes the likelihood of establishing a direct line of sight between client c and its associated UAV and is expressed as follows:

$$P_{\text{LOS}}(d_i) = \frac{1}{1 + a\exp\left(-b\left[\frac{180}{\pi}\arctan\left(\frac{h_U}{d_i}\right) - a\right]\right)}, \tag{2}$$

in this context, the environmental parameters a and b are constants, and their values are specified in [12]. The variable h_U denotes the height of the UAV and d_i represents the horizontal distance between client c and the center ground of the UAV. The likelihood of a non-line of sight (NLOS) scenario is expressed as $P_{\text{N}}(d_i) = 1 - P_{\text{L}}(d_i)$. Additionally, for both line of sight (LOS) and NLOS conditions, it is assumed that the fading follows the Nakagami-m distribution where all the power gains are modeled as independent and identically distributed (IID) gamma variables with shape parameters m_{L} and m_{N} respectively. Additionally, we assume a distance-dependent power-law path loss characterized by γ_L for LOS and γ_N for NLOS conditions. Furthermore, it is assumed that

the inherent noise N_0 at the base station follows a Gaussian distribution with variance σ^2.

To enhance the desired signal and reduce interference, we assume that both UAVs and devices utilize directional antennas. The antenna patterns are simplified using a discretized sectored gain model [17]. We denote the main lobe and side lobe by G_{uM} and G_{um}, respectively, for the UAV. Following the same notation, the main and side lobes for the client's devices are G_{cM} and G_{cm} respectively. Additionally, we assume that the link between the UAV and its associated client's device is perfectly aligned with $G_0 = G_{uM}G_{cM}$. However, the interfering links have uniformly distributed beam directions with gains G $\in G_{uM}G_{cM}, G_{uM}G_{cm}, G_{um}G_{cM}, G_{um}G_{cm}$ with probability $\mathbb{P} \in (\frac{\theta_u}{2\pi})(\frac{\theta_d}{2\pi}), (\frac{\theta_u}{2\pi})(1-\frac{\theta_d}{2\pi}), (1-\frac{\theta_u}{2\pi})(\frac{\theta_d}{2\pi}), (1-\frac{\theta_u}{2\pi})(1-\frac{\theta_d}{2\pi})$

To expedite the convergence speed of FL, the model selectively prioritizes high data rate clients during the initial rounds while temporarily excluding others. As the training progresses, the system gradually incorporates an increasing number of lower data rate clients to ensure broader participation over time. Adhering to this approach, clients with high data rates play a pivotal role in rapidly constructing a robust model, significantly reducing the overall training time. Therefore, the transmission rate in the t^{th} round is determined as $\log(1+\mu_t)$, where the transmission of the n^{th} client is deemed successful only if $\mathrm{SINR}_c^{(t)} > \mu_t$. A visual representation of this dynamic SINR training process is depicted in Fig 1.

C. Trust Model

Each client is assigned a trustworthiness score that is evaluated based on various critical factors. Noteworthy among these considerations is the historical engagement in FL processes and device type (whether it belongs to a trusted entity or a third-party client.

This score is determined by considering two main factors through a dual-faceted approach, as follows:

- **Security:** This criterion yields information about the authentication and security status of the client (e.g., security updates and patches).
- **Quality:** This criterion yields information on the level of the quality of the data and the capabilities of the devices for processing.

combining both metrics provides a unified metric known as the clients' trustworthiness measure, as cited in [8] and [13]. The trustworthiness of a client i is a crucial metric that would help in determining the best action to take for this client. In this system, we model it using a beta distribution with parameters α and β. The adoption of the Beta distribution is motivated by its flexibility and adaptability as a continuous distribution spanning from 0 to 1, offering a nuanced representation that can dynamically adjust to changes in the trustworthiness score.

The integration of both metrics results in a consolidated measure known as the clients' trustworthiness, as referenced in [8] and [13]. The trustworthiness of a client i denoted as χ_i, stands as a pivotal metric guiding decision-making for this client. Within this system, we represent this metric using the beta distribution, characterized by parameters α and β. The selection of the beta distribution is based on its versatility and adaptability, as it constitutes a continuous distribution that ranges from 0 to 1. This characteristic affords a nuanced representation that dynamically accommodates changes in the trustworthiness score.

The clients are categorized into three main sections. First, clients have trustworthiness scores greater than or equal to κ. These clients are classified as authenticated clients, denoted as $C_A \subset C$, where κ represents an upper threshold distinguishing reliable devices equipped with the latest security patches. Conversely, clients that have χ_i less than or equal to ρ are identified as hostile clients, forming the set $C_H \subset C$, with ρ serving as a low threshold indicative of potentially compromised devices or adversarial clients. Clients falling within the range $\kappa < \chi_i < \rho$ are categorized as risky clients, constituting the set $C_R \subset C$, and are flagged for potentially transmitting modified weights. This weight modification may arise from several factors like imprecise computations at the device level, a form of concealed (difficult to detect) model poisoning attack, as discussed in [14], or hosting a low-quality dataset. The deviation is assumed to be inversely proportional to the trustworthiness score χ_i and it is given by:

$$\boldsymbol{w}'_{i,t} = \boldsymbol{w}_{i,t} * \left(1 + \frac{(1-\chi_i)}{10}\right). \tag{3}$$

Adopting this modeling approach allows us to incorporate the influence of clients with varying trust scores, effectively capturing the consequences of dealing with risky clients. By applying this equation, we simulate the manipulative actions of clients deemed risky, addressing model imperfections introduced by clients with moderate trust scores.

It is important to emphasize that each client in set C exclusively falls into one of the trustworthiness categories, satisfying the conditions $C_A \cup C_R \cup\ C_H = U$ and $C_A \cap C_R \cap\ C_H = \phi$. Although clients in C_A are considered authenticated clients, they may lack sufficient data for FL training. As a solution, we suggest accommodating the risky clients C_R during the initial training phases. Subsequently, for the later stages of the learning process, we propose restricting the involvement of C_R to fine-tune the model.

III. Aerial Wireless Federated Safeguard in Heterogeneous Client Environments

Proof: See Appendix A. ∎

A. Success Uploading Model

To ensure a robust signal strength received by the UAV, the uplink SINR must exceed a predefined threshold, denoted as μ_t. If this criterion is not met, the signal quality falls below the required level for successful data decoding. However, omitting clients who do not meet the trustworthiness score requirements might lead to an uneven distribution of contributions to the global model's training process. To address this potential imbalance, we propose assigning a successful contribution distance-based weight. This weight amplifies the contribution

$$S_{i,t} = \sum_{n=1}^{m_L} (-1)^{n+1} \binom{m_L}{n} \exp\left(-N_0 \zeta_L\right) \mathcal{L}_{\mathcal{L}}\left(\zeta_L\right) \mathcal{L}_{\mathcal{N}}\left(\zeta_L\right) P_L(d_i) + \sum_{n=1}^{m_N} (-1)^{n+1} \binom{m_L}{n} \exp\left(-N_0 \zeta_N\right) \mathcal{L}_{\mathcal{L}}\left(\zeta_N\right) \mathcal{L}_{\mathcal{N}}\left(\zeta_N\right) P_N(d_i). \tag{4}$$

of devices located farther away due to their lesser participation in the communication rounds.

For a randomly selected client i ,the distance-dependent weighted factor is the reciprocal the success probability $S_{i,t} = (S_{i,t}|i = L)P_L(d_i) + (S_{i,t}|i = N)P_{NLos}(d_i)$. Let $\mathcal{L}_x(\cdot)$ be the Laplace transform of the interference probability density function and $\zeta_x = \frac{g_x(d_i^2+h^2)^{\gamma_x/2}\mu_t}{PG_0}$ for $x \in \{L, N\}$, then the transmission success probability $S_{i,t}$ can be represented as in (4) by averaging over the spatial and channel statistics. The Laplace transform of the interference in LOS links in the uplink, denoted as $\mathcal{L}_{\mathcal{L}}(\cdot)$, can be derived through systematic stochastic geometry analysis [15], and is given by

$$\mathcal{L}_x(s) = \exp\left\{-2\pi\lambda \sum_{q=1}^{4} \mathbb{P}_q \int_0^{\infty} \left(1 - \exp\{-\pi\lambda z^2\}\right) \left(1 - \left(1 + \frac{sPG_q(z^2+h^2)^{-\gamma_x/2}}{m_x}\right)^{-m_x}\right) z \quad dz\right\}. \tag{5}$$

Following the same notation, the Laplace transform of the interference in NLOS links in the uplink, denoted as $\mathcal{L}_{\mathcal{NLOS}}(\cdot)$ has a similar expression while replacing m_L with m_n and $P_L(d_i)$ with $P_n(d_i)$. It is worth noting that (4) is derived using standard stochastic geometry analysis by following the same steps as in [2], [5], [7], which is detailed in the appendix.

The expression in (4) defines the transmission success probability ($S_{i,t}$) for client i in round t. This probability is based on the likelihood that client i achieves an SINR exceeding the predefined threshold μ_t. Upon receiving client weights, the UAV unbiases the contribution of successful clients by multiplying their contributions with a factor of $\frac{1}{S_{i,t}}$. This unbias factor aims to compensate for the different contributions of clients due to their diverse success transmission probabilities, promoting a more balanced and equitable aggregation process.

B. Dynamic SINR *Thresholds and Trustworthiness Integration: Algorithm Details*

In **Algorithm 1** the detailed behavior of the system is represented. The system groups both risky and authenticated clients $C_A \cup C_R$ into a subgroup C'. This inclusion of clients proves particularly effective in the context of non-IID datasets, where diverse devices with distinct datasets contribute positively to system enhancement. After that, another set $C'' \in C'$ represents the set of clients who have the required SINR

$$\boldsymbol{g_{t+1}} \leftarrow \boldsymbol{g_t} + \frac{1}{C''} \sum_{n=1}^{C'} \frac{\mathbb{1}\{\text{SINR}_i^{(t)} > \mu_t\}}{S_{i,t}} (\boldsymbol{w}'_{i,t} - \boldsymbol{g_t}), \tag{6}$$

in the given equation, the indicator fuction $\mathbb{1}\{\cdot\}$ signifies the inclusion of the weights corresponding to the client i only when it's (SINR) surpasses the predefined threshold μ_t. This selective consideration of clients ensures that only those with sufficiently favorable SINR conditions contribute to the global model update.

Algorithm 1: Aerial Wireless Federated Safeguard

Data: $T, \mu_t, \boldsymbol{w}_0, E, \eta, \Omega$
Result: $\boldsymbol{g}_t$
Initialization;
$\boldsymbol{g}_0 \leftarrow$ First Global Model;
$\boldsymbol{Q} \leftarrow$ List of len(T) elements ;
$C' \leftarrow C_A \cup C_R$;
$t \leftarrow 0$;
while $t < T$ **do**
  $n \leftarrow 1$;
  while $n \leq C'$ **do** *for each client*
    $\boldsymbol{w}_{i,t}^{(0)} \leftarrow \boldsymbol{g_t}$;
    $x \leftarrow 0$;
    while $x \leq (E-1)$ **do**
      $\boldsymbol{w}_{i,t}^{(x+1)} \leftarrow \boldsymbol{w}_{i,t}^{(x)} - \gamma \boldsymbol{g}_i(\boldsymbol{w}_{i,t}^{(x)})$;
    end
    $\boldsymbol{w}'_{x,t} \leftarrow \boldsymbol{w}_{x,t}^{(E)} \cdot \left(1 + \frac{(1-\gamma_x)}{10}\right)$;
    transmit $\boldsymbol{w}'_{i,t}$;
    if $\text{SINR}_i^{(t)} > \mu_t$ **then**
      Add client i to C'';
    end
    $n \leftarrow n + 1$;
  end
  $\boldsymbol{g}_{t+1} \leftarrow \boldsymbol{g}_t + \frac{1}{C''} \sum_{n=1}^{C''} \frac{1}{S_{i,t}} (\boldsymbol{w}'_{i,t} - \boldsymbol{g}_t)$;
  $\boldsymbol{Q}[t] \leftarrow$ Assess the accuracy ($\boldsymbol{g_{t+1}}$);
  if $\boldsymbol{Q}[t]$ $<$ *the previous* μ *values* **then**
    $C' \leftarrow C_A$;
  end
  $t \leftarrow t + 1$;
end
return Result;

IV. Numerical Results

We consider wireless network, spanning an expansive area of 2500 x 2500 km^2, the UAV density (λ) is considered to be $50/km^2$, positioned at an elevation h of 45 m. Each UAV is equipped with 30 Resource Blocks (RBs), enabling it to effectively serve up to 30 devices. These devices transmit with a power (P) of 10 dBm. The path loss exponent for LOS γ_L is set at 2.5, while for NLOS γ_N is configured to 4. Furthermore,

the main lobe beamwidth (θ) for both the UAV and clients' devices is designated as 40°. In addition, a main gain lobe of 3.162 is assumed for both the UAV and the devices, accompanied by a side gain lobe of 1. These specifications provide a comprehensive overview of the wireless network parameters, offering insights into the configuration and capabilities of the UAVs and devices within the designated geographical expanse. We chose the parameters of the trustworthiness metric to get a mean of "90%" and "80%". To achieve this, we use $\alpha = 8, \beta = 1$ and $\alpha = 8, \beta = 2$ respectively. We used Convolutional Neural Network (CNN) architecture for our model with stochastic gradient descent (SGD) as an optimizer with two epochs and 0.5 momentum. We tested the proposed model on classifying the MNIST dataset [16]. We set κ and ρ to be 0.9 and 0.3, respectively. The dynamic threshold associated with the SINR ranges from 10 dB to 1 dB with a step size of 0.25.

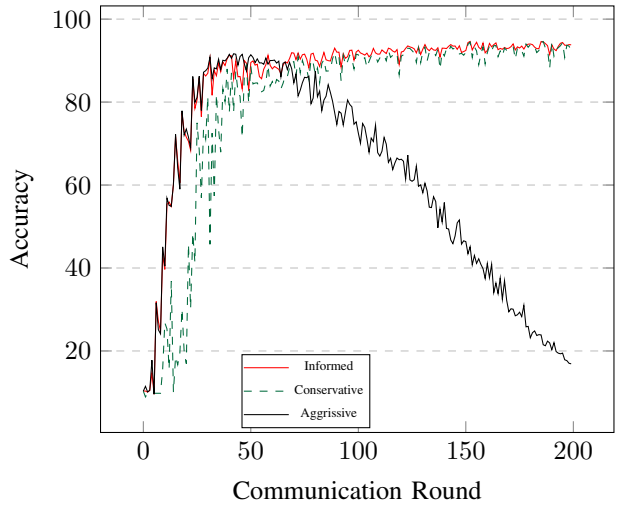

Fig. 2. Global Accuracy vs Communication Rounds with mean = 0.90

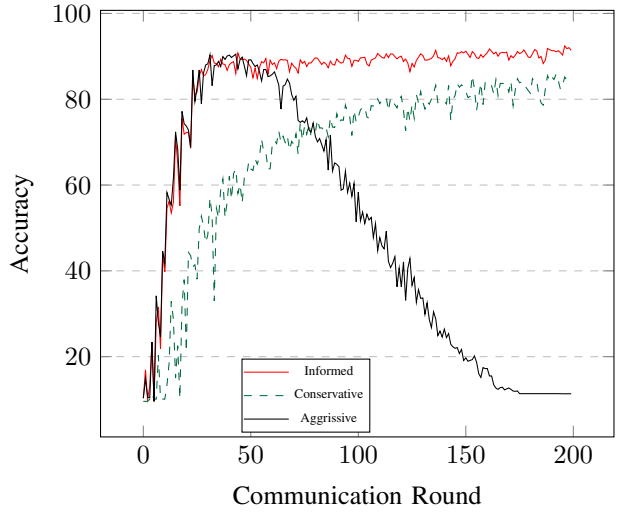

Fig. 3. Global Accuracy vs Communication Rounds with mean = 0.80

In these figures, three cases are compared:

- **Conservative:** accommodating only authenticated clients C_A in the aggregation process
- **Aggressive**: accommodating authenticated C_A and risky C_R clients for the entire learning process
- **Informed:** In initial rounds, the system accommodates authenticated C_A and risky C_R clients. Then, when the accuracy starts to degrade, it discards the risky client from the learning process.

Several observations come to light from Figs. 2 and 3. First, while aggressively accommodating risky clients C_R accelerates the convergence speed of the model, and it reaches a critical threshold beyond which it detrimentally impacts the overall learning model. This is because of their accumulated model drift over many iterations. Furthermore, it is imperative to note that adhering strictly to a conservative approach may not always be the optimal solution, particularly when there is a scarcity of participating clients in a non-IID dataset environment. On the contrary, accommodating risky clients initially and excluding them in subsequent rounds would result in improved accuracy and faster convergence, particularly in environments with lower trustworthiness.

Derived from the numerical findings, several key observations come to light. Initially, it becomes evident that solely embracing a conservative approach proves inadequate concerning both the convergence rate and the peak accuracy attained. This inadequacy is particularly pronounced in networks characterized by lower trustworthiness, as illustrated in Fig. 3 and Fig. 2.

V. Conclusion

We examined the nuanced interplay between accommodating and avoiding risky clients in the context of Federated Learning (FL). Additionally, we observed that the involvement of risky clients significantly contributes to the construction of the initial model, particularly in scenarios where authenticated clients are limited. Expanding on this observation, we devised a novel approach that dynamically shifts from relying on risky clients to exclusively engaging authenticated clients as the learning process progresses. This adaptive strategy allows us to capitalize on the benefits of involving risky clients in the initial model development, while simultaneously ensuring a more secure and dependable learning environment in subsequent stages.

VI. Appendix

This appendix details the proof of the success probability in (4). The success probability for a LoS client is given by

$$(S_{i,t}(\mu_t, d_i)|i = L) = \mathbb{P}(\text{SINR}_i^{(t)} > \mu_t)$$
$$(S_{i,t}(\mu_t, d_i)|i = L) = \mathbb{P}\left(\frac{PG_0 h_0^2 (d_i^2 + h^2)^{-\gamma_L/2}}{N_0 + I_{Los} + I_{NLos}} > \mu_t\right)$$

since h_0^2 follows a gamma distribution, using Alzer's inequality in [18] gives us an approximate of

$$\approx \sum_{n=1}^{m_L} (-1)^{n+1} \binom{m_L}{n} \cdot \mathbb{E}_{Iagg}\left[\exp\left(\frac{-S_L n \mu_t \left(N_0 + I_L + I_N\right)}{P G_0 (d_i^2 + h^2)^{-\gamma_L/2}}\right)\right],$$

where $S_L = m_L (m_L!)^{\frac{-1}{m_L}}$

$$= \sum_{n=1}^{m_L} \left((-1)^{n+1} \binom{m_L}{n} \exp\left(\frac{-\mu_t N_0 S_L n (d_i^2 + h^2)^{\gamma_L/2}}{P G_0}\right) \cdot \mathcal{L}_{\mathcal{L}}\left(\frac{S_L n (d_i^2 + h^2)^{\gamma_L/2} \mu_t}{P G_0}\right) \cdot \mathcal{L}_{\mathcal{N}}\left(\frac{S_L n (d_i^2 + h^2)^{\gamma_L/2} \mu_t}{P G_0}\right)\right).$$

Following the same steps for a NLOS client, we replace m_L with m_N and S_L with S_N have

$$(S_{i,t}(\mu_t, d_i)|i = N) = \frac{\sum_{n=1}^{m_i} (-1)^{n+1} \binom{m_L}{n}}{\exp\left(\frac{\mu_t N_0 S_N n (d_i^2 + h^2)^{\gamma_L/2}}{P G_0}\right)} \cdot \mathcal{L}_L\left(\frac{S_N (d_i^2 + h^2)^{\gamma_N/2} \mu_t}{P G_0}\right) \cdot \mathcal{L}_N\left(\frac{S_N (d_i^2 + h^2)^{\gamma_N/2} \mu_t}{P G_0}\right)$$

Now, considering the law of total probability,

$$(S_{i,t}) = (S_{i,t}|i = L) P_{Los}(d_i) + (S_{i,t}|i = N) P_N(d_i)$$

$$= \frac{\sum_{n=1}^{m_L} (-1)^{n+1} \binom{m_L}{n}}{\exp\left(\frac{\mu_t N_0 S_L n (d_i^2 + h^2)^{\gamma_L/2}}{P G_0}\right)} \mathcal{L}_L\left(\frac{S_L (d_i^2 + h^2)^{\gamma_L/2} \mu_t}{P G_0}\right) \cdot \mathcal{L}_N\left(\frac{S_L (d_i^2 + h^2)^{\gamma_L/2} \mu_t}{P G_0}\right) P_L(d_i) +$$

$$\frac{\sum_{n=1}^{m_i} (-1)^{n+1} \binom{m_L}{n}}{\exp\left(\frac{\mu_t N_0 S_N n (d_i^2 + h^2)^{\gamma_L/2}}{P G_0}\right)} \cdot \mathcal{L}_L\left(\frac{S_N (d_i^2 + h^2)^{\gamma_N/2} \mu_t}{P G_0}\right) \cdot \mathcal{L}_N\left(\frac{S_N (d_i^2 + h^2)^{\gamma_N/2} \mu_t}{P G_0}\right) P_N(d_i),$$

which proves (4).

REFERENCES

[1] T. Li, A. K. Sahu, A. Talwalkar, and V. Smith, "Federated learning: Challenges, methods, and future directions," IEEE Signal Processing Magazine, vol. 37, no. 3, pp. 50–60, 2020.

[2] Y. Nabil, H. ElSawy, S. Al-Dharrab, H. Attia and H. Mostafa, "Ultra-Reliable Device-Centric Uplink Communications in Airborne Networks: A Spatiotemporal Analysis," in *IEEE Transactions on Vehicular Technology*, vol. 72, no. 7, pp. 9484-9499, July 2023, doi: 10.1109/TVT.2023.3250757. ,

[3] McMahan, B., Moore, E., Ramage, D., Hampson, S. amp; Arcas, B.A.y.. (2017). Communication-Efficient Learning of Deep Networks from Decentralized Data. Proceedings of the 20th International Conference on Artificial Intelligence and Statistics, in Proceedings of Machine Learning Research 54:1273-1282 Available from https://proceedings.mlr.press/v54/mcmahan17a.html.

[4] M. Salehi and E. Hossain, "Federated Learning in Unreliable and Resource-Constrained Cellular Wireless Networks," in IEEE Transactions on Communications, vol. 69, no. 8, pp. 5136-5151, Aug. 2021, doi: 10.1109/TCOMM.2021.3081746.

[5] R. Zhagypar, N. Kouzayha, H. ElSawy, H. Dahrouj and T. Y. Al-Naffouri, "Characterization of the Global Bias Problem in Aerial Federated Learning," in IEEE Wireless Communications Letters, vol. 12, no. 8, pp. 1339-1343, Aug. 2023, doi: 10.1109/LWC.2023.3273318.

[6] H. Xia, Y. Li, C. Liu and Y. Zhu, "Stochastic Client Scheduling with Dynamic SINR Thresholds for Fast Federated Learning," 2022 IEEE/CIC International Conference on Communications in China (ICCC), Sanshui, Foshan, China, 2022, pp. 632-637, doi: 10.1109/ICCC55456.2022.9880839.

[7] M. Ads, H. ElSawy and H.S. Hassanein, "Risk-Aware Accelerated Wireless Federated Learning with Heterogeneous Clients," *ICC 2024 - IEEE International Conference on Communications*, Denver, USA, Accepted

[8] Gholami, A., Torkzaban, N. and Baras, J.S., 2022, January. Trusted Decentralized Federated Learning. In 2022 IEEE 19th Annual Consumer Communications and Networking Conference (CCNC) (pp. 1-6). IEEE.

[9] Wahab, O.A., Rjoub, G., Bentahar, J. and Cohen, R., 2022. Federated against the cold: A trust-based federated learning approach to counter the cold start problem in recommendation systems. Information Sciences, 601, pp.189-206.

[10] L. Yi, X. Shi, W. Wang, G. Wang and X. Liu, "FedRRA: Reputation-Aware Robust Federated Learning against Poisoning Attacks," 2023 International Joint Conference on Neural Networks (IJCNN), Gold Coast, Australia, 2023, pp. 1-8, doi: 10.1109/IJCNN54540.2023.10191556.

[11] X. Cao, M. Fang, J. Liu, and N. Z. Gong, "FLTrust: Byzantine-robust Federated Learning via Trust Bootstrapping," in 28th Annual Network and Distributed System Security Symposium (NDSS), 2021, virtually, February 21-25, 2021. [Online]. Available: [https://www.ndss-symposium.org/ndss-paper/fltrust-byzantine-robust-federated-learning-via-trust-bootstrapping/].

[12] A. Al-Hourani, S. Kandeepan, and S. Lardner, "Optimal LAP altitude for maximum coverage," IEEE Wireless Commun. Lett., vol. 3, no. 6, pp. 569–572, Dec. 2014

[13] Kumar, R. and Goyal, R., 2022. Performance based Risk driven Trust (PRTrust): On modeling of secured service sharing in peer-to-peer federated cloud. Computer Communications, 183, pp.136-160.

[14] N. Tezuka, H. Ochiai, Y. Sun and H. Esaki, "Resilience of Wireless Ad Hoc Federated Learning against Model Poisoning Attacks," 2022 IEEE 4th International Conference on Trust, Privacy and Security in Intelligent Systems, and Applications (TPS-ISA), Atlanta, GA, USA, 2022, pp. 168-177, doi: 10.1109/TPS-ISA56441.2022.00030.

[15] H. ElSawy, A. Sultan-Salem, M. -S. Alouini and M. Z. Win, "Modeling and Analysis of Cellular Networks Using Stochastic Geometry: A Tutorial," in IEEE Communications Surveys & Tutorials, vol. 19, no. 1, pp. 167-203, Firstquarter 2017, doi: 10.1109/COMST.2016.2624939.

[16] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," Proceedings of the IEEE, vol. 86,

[17] J. G. Andrews et al., "Modeling and analyzing millimeter wave cellular systems," IEEE Trans. Commun., vol. 65, no. 1, pp. 403–430, Jan. 2017.

[18] H. Alzer, "On some inequalities for the incomplete gamma function," Math. Computation, vol. 66, no. 218, pp. 771–778, 1997.