Channel-Aware Federated Analytics in B5G/6G Networks: Dynamic Power Allocation with NS-3 5G-LENA

Xavier P. Sebastião [1,2], Renan R. de Oliveira[1,3], Waldir Moreira[4], Antonio Oliveira-Jr[1,4]

[1]Instituto de Informática – Universidade Federal de Goiás (UFG) – Goiânia – Brazil

[2]Faculdade de Ciências Agrárias – Universidade Zambeze (UZ) – Angónia – Mozambique

[3]Instituto Federal de Goiás (IFG) – Goiânia – Brazil

[4]Fraunhofer Portugal AICOS – Porto – Portugal

xavierpaulino@discente.ufg.br, renan.rodrigues@ifg.edu.br,

waldir.junior@fraunhofer.pt, antoniojr@ufg.br

Abstract. *Federated Analytics (FA) is an approach for preserving security and privacy by implementing collaborative analysis of data from distributed devices without sharing raw data. However, when FA operates over wireless transmission, challenges such as interference, signal degradation, and network congestion may arise. These factors can make the wireless transmission unreliable, introducing delays and causing corruption in responses and updates received at the central server, thereby impacting the quality of the final aggregated FA results. This work proposes an integrated framework to simulate FA in real 5G conditions using NS-3 5G-LENA. It applies two algorithms: a channel-aware power allocation algorithm for optimized transmission power allocation and a synchronous FA-5GLENA algorithm for the FA and 5G-LENA integration. Simulation results show the channel-aware algorithm outperforms uniform and random power allocation in both network and FA performance. FA accuracy reached 93.17 %, precision 93.31 %, and recall 93.09 %, statistically significantly higher than uniform (55.96 %, 56.02 %, 55.90 %) and random (42 %, 42.02 %, 41.96 %) allocation. These findings demonstrate the algorithm's superiority in enhancing FA within 5G networks.*

1. Introduction

The global interconnectivity and the growing number of connected devices across industries have generated an unprecedented volume of data, driving demand for advanced data science techniques to extract valuable insights to support decision-making processes. However, centralizing this data for analysis is not always feasible due to privacy, security, and communication efficiency considerations [Elkordy et al. 2023, Wang et al. 2024]. In this context, Federated Analytics (FA) emerges as an innovative solution, allowing the collaborative analysis of data distributed between multiple entities without the need to share the raw data. This approach not only meets stringent data privacy and regulatory requirements, but also offers an efficient way to address the latency, energy and costs associated with transferring large volumes of data, thereby maximizing the utility of the data while minimizing the risks associated with its handling and transfer.

From a general perspective, FA can be defined as a scenario for data analysis where a questioner wants to respond to a data analysis query through the collaboration of multiple proprietary data devices that have their local raw data. Raw data is not exchanged or transmitted, but instead, intermediate query responses intended for aggregation at the questioner are transferred to answer the intended query. The term FA was first coined by Google in 2020 to represent "collaborative data science without data collection" [Elkordy et al. 2023]. It was first explored in support of Federated Learning (FL) as a way for Google engineers to evaluate the quality of learned Machine Learning (ML) models against real-world data.

While FA deals with a wide range of analytical issues, FL focuses on collaborative model training, allowing local devices or servers to contribute to the learning of a common model without exposing the underlying data. Although both share the principle of minimizing data sharing, they differ in application and complexity. FL involves iterative and computationally intensive processes to update machine learning models, while FA can involve anything from simple aggregations to complex analytical operations [Kairouz et al. 2021].

FA use cases span a variety of domains that significantly benefit from the ability to perform data analysis without compromising the privacy of participants [Elkordy et al. 2023]. In technology, it allows you to evaluate the accuracy of prediction models like those from Gboard, using anonymous typing data. In healthcare, it facilitates diagnosis and research with large volumes of data, without exposing sensitive patient information, supporting studies on treatments and symptoms. In marketing, FA helps advertisers understand the effectiveness of their ads through anonymous viewing data, optimizing advertising strategies without compromising users' privacy.

However, FA faces significant challenges when implemented in wireless networks. In this context, limited resources, interference, signal degradation, and network congestion can result in delays in data transmission and occasionally packet loss [Chen et al. 2021]. In this case, the unreliable nature of wireless transmission channels can compromise the integrity of the transmissions forwarded to the questioner, which harms the construction of the response based on data distributed across different devices.

This work emphasizes the promoting of FA as a critical component in the transition to B5G/6G networks, addressing the challenges of unreliable wireless transmission channels. We highlight the development of algorithms that integrate FA with the NS-3 5G-LENA simulator to optimize FA performance under real-world 5G network conditions. The work focuses on effectively managing transmission power to maximize performance in future networks. We propose an integrated framework and specific algorithms for realistic FA simulations, capturing both contemporary and emerging network conditions. This systematic approach allows a detailed assessment of how changing network conditions, such as interference and signal degradation, impact FA performance. Finally, we validate the results through an application in collaborative quality assessment of trained ML models on classification problems, ensuring the relevance and applicability of the developed solutions.

The principal contributions of this article are outlined as follows: (*i*) **Network Adaptive FA Algorithm:** We present the implementation of FA algorithms specifically adapted to the characteristics of 5G networks; (*ii*) **FA with NS-3 Simulator:** We developed a framework that integrates FA with the NS-3 5G-LENA simulator. This facilitates realistic and scalable simulations of FA under diverse network conditions; (*iii*) **Performance Evaluation:** We validate our algorithm through an application in the domain of collaborative quality assessment of trained ML models on classification problems; (*iv*) **Validation and Documentation:** Our evaluation included analysis that evaluates how various network conditions, such as packet loss, latency, and throughput, affect FA accuracy, precision and recall, demonstrating their effectiveness in tackling challenges related to 5G wireless networks; (*v*) **Source Code Sharing:** The source code [1] has been made accessible to enable replication and validation of results, serving as a valuable resource for future researchers and developers in the FA domain for B5G/6G networks. Available at https://github.com/LABORA-INF-UFG/fa-5glena.

The rest of the paper is organized as follows. Section 2 discusses related works. Section 3 presents the system model and network model for FA in 5G. Section 4 describes the FA and 5G-LENA integrated algorithms. Section 5 provides configuration parameters and simulation results. Finally, Section 6 presents the final considerations and indicates directions for future work.

2. Related Work

Existing literature has demonstrated the potential of FA in improving data privacy, reducing communication overhead, and improving scalability across distributed networks. Despite this success, specific optimizations projected for B5G/6G networks remain an active research area. The study by [Zhao et al. 2022] introduced a semi-hierarchical FA framework that integrates FL with edge computing architectures. It uses multiple edge servers to aggregate updates from IoT devices and combines learned model weights without relying on the cloud or a central server. Like our project, their work includes an evaluation of network conditions on FA performance through packet loss analysis. However, it differs in its focus on edge computing for IoT, and our work incorporates more network metrics in the analysis, such as throughput, latency, delivery ratio and energy consumption.

The work done by [Shi et al. 2022] proposed a "Federated Anomaly Analytics enhanced Distributed Learning" framework that allows clients and the server to collaborate in analyzing anomalies, by identifying and eliminating all potential malicious trained local models, and aggregating the remaining models to generate the global model. This work illustrates the application of FA on anomaly detection, while our work explores the application of FA in future generation networks (B5G/6G). Differently, their study does not include the analysis of network impact on FA performance.

Applying FA techniques, [Yue et al. 2024] developed a "Federated Data Analysis (FDA)" approach to improve the statistical linear regression model. Their approach is based on hierarchical modeling, allowing the combination of information across multiple groups, thereby promoting collaborative analysis. In our work, we also apply statistical methods such as the Bayesian aggregation method. However, again their study does not include the analysis of network impact on FA performance.

The study from [Toka et al. 2023] focuses on the coordination of vehicles in autonomous driving, emphasizing the importance of communication and reducing latency in

[1] Available at https://github.com/LABORA-INF-UFG/fa-5glena

data collection, processing, and sharing. It applies a FA system that protects data privacy while keeping high-definition updated maps through crowd-sourcing. As our work, their project is concentrated on optimizing latency and information delivery for 5G networks. Besides latency, our work focuses on optimizing more metrics for 5G networks to generally minimize the impact on FA performance. Also, their project uses both mobile and fixed nodes, but our work only uses fixed nodes.

Designed for decentralized data analytics in connected vehicles, [Zhao et al. 2024] extended the concept of FL to support decentralized model training directly on vehicles, which helps eliminate the need for a centralized server. To improve communication efficiency, they introduce a federated regularized nonlinear acceleration-based local training method aimed at minimizing communication rounds. Related to our work, they concentrate their study on communication efficiency using ML techniques. This work also studies the communication efficiency but on a different perspective, which is network.

While the majority of these studies discuss the applications of FA, only a few address the impacts of network conditions. None of them evaluate in detail how the unreliable nature of wireless transmission medium affects FA. They neither integrate NS-3 for simulations nor perform network optimization for FA. Only one work attempts to apply network latency optimization on FA. They mainly focus on machine learning and the application of FA itself. It is important to note that the low quality of wireless networks can significantly affect both communication and FA performance.

Table 1. Comparison of this proposal and the related works

Work	Intervention	Metric	Contribution	Network Analysis	Integration with NS-3 or 5G-LENA
[Zhao et al. 2024]	Communication for connected vehicles	FA communication cost	FA framework	✗	✗
[Yue et al. 2024]	Federated hierarchical linear model structures	FA efficiency	Federated data analysis method	✗	✗
[Toka et al. 2023]	Communication for 5G-vehicular networks	Latency	Analytical model and optimization algorithm	✓	✗
[Zhao et al. 2022]	Communication for IoT	• FA convergence • Network packet loss	FA framework	✓	✗
[Shi et al. 2022]	Local models anomaly detection and elimination	• FA accuracy, robustness, efficiency	FA framework	✗	✗
This proposal	Communication for B5G/6G networks	• FA accuracy, precision, recall • Throughput, latency, packet loss, delivery ratio, energy consumed	FA framework and optimization algorithm	✓	✓

This proposal stands out as the most comprehensive project integrating the network on FA. Unlike related studies, which are primarily machine learning-focused, this work incorporates network performance evaluation as a core component for FA optimization. Besides using a different network technology (NS-3 with 5G-LENA), it considers a broader set of metrics, including throughput, latency, delivery ratio, and energy consumption. These metrics provide a clearer picture of network behavior from multiple perspectives, to ensure that network quality is maximized. This guarantees the efficient transmission of client responses to the server, which significantly contributes to FA performance optimization. See Table 1, for a detailed comparison between this work and related studies.

3. System Model

3.1. Federated Analytics

The central figure in FA is the querier, often referred to as the questioner, who seeks to obtain insights or answers from data that is distributed across various clients, also known as parties. Each client holds a segment of the overall data, termed their local dataset. FA essentially operates as a collaborative data analysis framework where the querier aims to resolve a data analytics query through the cooperation of multiple data owners who manage their own local raw data. Rather than exchanging or transmitting raw data, intermediate responses to queries are shared. These responses are subsequently aggregated by the querier to provide a comprehensive answer to the posed question. From this generalized idea, by [Elkordy et al. 2023], if we consider $D = \{D_i\}_{i=1}^{N}$ as the set of local datasets from each client, the goal of FA will be for a central querier to answer the query Q, given by $Q(D) = F_\omega(D_1, D_2, \ldots, D_N)$ where F_ω is the parameterized function on the data describing the target query.

3.2. Network Model for FA in 5G

Consider a 5G network composed of a Base Station (gNB) connected directly to a central FA server, as proposed in the *Multi-Access Edge Computing* (MEC) approach [ETSI 2022]. Suppose the existence of a set $S = \{i_1, i_2, ..., i_N\}$ with N active 5G devices before starting an analysis session. Each client in the network has a dataset $\mathcal{D}_i$ sample stored on their respective local devices. These devices are connected to the gNB via a 5G connection and are capable of collecting data and performing pre-processing necessary for specific federated analysis tasks.

Instead of transmitting raw data to the central server, devices perform local calculations to respond to specific queries issued by the FA server. Intermediate responses which may include summary statistics are transmitted to the gNBs. The aggregation of these intermediate responses is performed by the server to compile a consolidated analysis or a final result, maintaining the confidentiality of the original data from each device. To reinforce security and privacy, techniques such as Differential Privacy (DP) or Secure Multiparty Computing (MPC) can be integrated into the FA model for 5G networks. This arrangement minimizes the need for bandwidth for data transmission and increases privacy and security, taking advantage of the high capacity and low latency of 5G networks.

3.3. FA Application Domain

We validate our results through an application in the domain of ML model evaluation on classification problems. This motivation arises from the growing interest in FA, particularly in the collaborative assessment of the quality of ML models in relation to data from practical applications. This data, collected from real everyday situations, offers a more reliable and complex representation of the conditions and variability that the models will face in real use.

As defined by [Elkordy et al. 2023], suppose we have a pre-trained ML classification model defined by ω. Then, in FA, a typical query to evaluate the model's accuracy on distributed datasets can be formulated as

$$Q_\omega(D) = Acc\left(\omega; \{D_1, D_2, \ldots, D_N\}\right) = \sum_{i=1}^{N} \frac{|D_i|}{\sum_{i=1}^{N} |D_i|} Acc(\omega; D_i), \tag{1}$$

where the answer to the query is determined by calculating the weighted average of each client's local test accuracy, denoted as $Acc\,(\omega; D_i)$. The term $\sum_{i=1}^{N} |D_i|$ serves as the normalizer. To find the local accuracy, each client uses its model to evaluate its own labeled dataset and locally computes the ratio of correct classifications. Following this idea, in this article we use a weighted averaging algorithm to calculate the aggregated accuracy and extend the method to aggregation of precision and recall, all with incorporation of reliability score (confidence).

3.4. FA Integration with 5G-LENA Module

We have developed an integrated FA, NS-3 and 5G-LENA framework to analyze how different network conditions such as latency, packet loss, throughput, and energy consumption affect final response quality and FA performance in the context of real-world network conditions of 5G networks. NS-3 is an open source discrete event network simulator, licensed under the GNU GPLv2 [Zárate Ceballos et al. 2021]. It is a powerful research tool designed to model network elements, simulate complex network behaviors, and analyze the performance of network protocols in a controlled environment [Hapanchak and Costa 2022]. 5G-LENA is a module of NS-3 that simulates 5G New Radio (NR) networks, which allows the simulation and evaluation of the potential of 5G solutions [Larrañaga et al. 2023].

The integration was made through a shared memory technique, allowing the FA to access and use the network results. Both the server and clients operate synchronously in this setup. Each simulation round starts with the FA server connecting to the NS-3 5G-LENA simulator via gNB, where it sends a network simulation request command (netrun) including initial parameters such as number and list of clients, and the query. The server oversees the start, progression, and conclusion of the simulation for that round. After dispatching the initial parameters, the server awaits responses. Once the network request command is received, the 5G-LENA simulator performs the simulation, outputting network statistics like latency, packet loss, delivery ratio, throughput, and energy consumption for each client. To signal the end of the network simulation, 5G-LENA sends a command (netend) for the FA server. The network statistics are then used by the FA server in conjunction with FA algorithms to select clients for subsequent rounds and to aggregate results at the end of the simulation. See the system architecture given in Figure 1.

4. The Channel-Aware Transmission Power Allocation Algorithm

We developed an adaptive optimization algorithm that efficiently manages and allocates transmission power according to channel conditions. This algorithm dynamically calculates the minimum transmission power required for a UE to achieve a delivery ration DR $\geq$ 99 $\%$. It begins by estimating the initial power based on network parameters such as distance, frequency, bandwidth, and path loss (considering both Line-of-Sight and Non-Line-of-Sight conditions) of 3gpp UMi models [38.901 2017]. By applying feedback in a closed-loop technique, iteratively adjusts the power, calculating the DR using SNR, BER, and PER. If the DR is below 99 $\%$, it increases the power; if the DR exceeds 99 $\%$, it reduces the power. The iterative process continues until the DR reaches 99 $\%$ (convergence) or the maximum number of iterations is reached. The final output is the minimum power required to meet DR $\geq$ 99 $\%$, while ensuring efficient power usage within limits. See algorithm 1 for detailed operation.

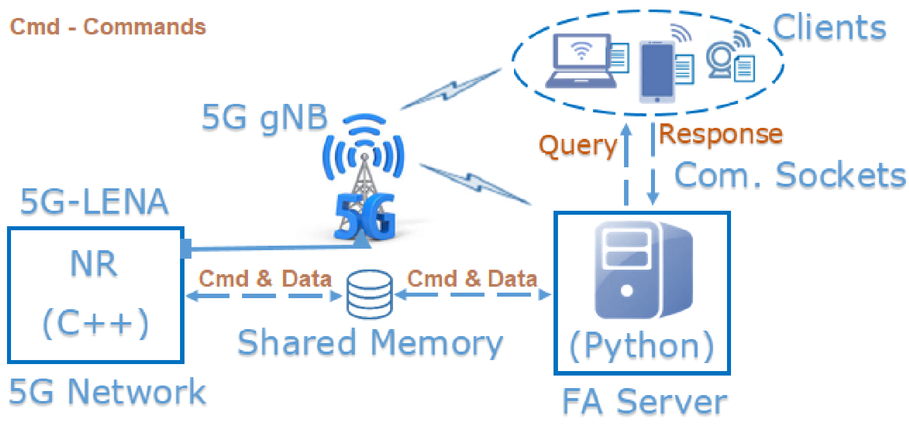

Figure 1. FA Integration with 5G-LENA Module

4.1. Objective Function for Power Allocation by Channel-Aware Algorithm

The main objective of this project is to improve communication between FA clients and the server in a 5G network by optimizing transmission power. The algorithm aims to dynamically allocate power for each UE to balance an increased delivery ratio and throughput with reduced latency and energy consumption. To achieve these goals based on transmission power as the main variable, the objective function is formulated as

$$min \sum_{i=1}^{N} \left(P_i^{(t)} \right) \tag{2}$$

Subject to

- $0 \ \leq \ P_i^{(t)} \ \leq \ 30\, dBm, \ \ \forall i \in \{1, 2, 3, ..., N\}$
- $\mathfrak{T}_i^{(t)}(P_i^{(t)}) \ \geq \ 100\, Mbps, \ \ \forall i$
- $\mathcal{L}_i^{(t)}(P_i^{(t)}) \ < \ 1\, ms, \ \ \forall i$
- $\mathcal{DR}_i^{(t)}(P_i^{(t)}) \ \geq \ 99\,\%, \ \ \forall i$

Where

$P_i^{(t)}$ is the transmission power for the $i-th$ UE at iteration t. $\mathfrak{T}_i^{(t)}(P_i^{(t)})$ is the average end-to-end throughput for received packets from the $i-th$ UE, as a function of transmission power, measured in 5G-LENA. $\mathcal{L}_i^{(t)}(P_i^{(t)})$ is the average end-to-end latency for received packets from the $i-th$ UE, as a function of transmission power, measured in 5G-LENA. $DR(P_i^{(t)})$ is the delivery ratio for the $i-th$ UE, as a function of transmission power, calculated by the formula $DR_i^{(t)}(P_i^{(t)}) = 1 - PER_i^{(t)}(P_i^{(t)})$.

4.2. Synchronous FA-5GLENA Integrated Algorithm

The algorithm first simulates the NS-3 5G-LENA network to collect performance statistics (packet loss, throughput, latency, delivery ratio, and energy consumption) for 20 clients. Then, the statistics are utilized in the FA process, where the server selects clients with a delivery ratio DR $\geq$ 90 % to participate in the round. These selected clients train the model, then make classifications on MNIST based on received query, and send responses along with performance metrics to the server. Non-selected clients send error

Algorithm 1 Channel-Aware Transmission Power Allocation Algorithm

Initialize for Each UE i:
Calculate path loss models (LoS and NLoS) $L_i^{(t)}$ based on 3gpp UMi loss models.
Calculate receiver sensitivity using noise power and target SNR.
Initial Transmit Power Calculation for each UE i:
Compute the initial transmit power using receiver sensitivity, path loss, and margin.
Main Iteration t for Each UE i:
for each UE i **do**
    Set $P_i^{(t)}$ to the initial calculated value.
    While Loop: Continue adjusting $P_i^{(t)}$ until convergence or maximum iterations.
    while $P_i^{(t)} >$ minimum transmit power and iteration $t <$ maximum **do**
        Calculate $DR_i^{(t)}$ based on $P_i^{(t)}$ and channel conditions.
        if $DR_i^{(t)} < 99$ % **then**
            Increase $P_i^{(t)}$ to improve $DR_i^{(t)}$.
        else
            Decrease $P_i^{(t)}$ to obtain minimum satisfying $DR_i^{(t)} \geq 99$ %.
        end if
        Adjust the step size as the $DR_i^{(t)}$ approaches 99 %, after feedback.
        if $DR_i^{(t)}$ is sufficiently close to 99 % **then**
            Exit loop early due to convergence.
        end if
        Increment the iteration t counter.
    end while
    Final Transmit Power $P_i^{(t)}$ for UE i: Return the calculated $P_i^{(t)}$ for UE i.
end for
Return: $P_i^{(t)}$ for UE i.

responses. The server aggregates all responses and determines the final answer to the query after each round. Aggregation is done with two methods, Bayesian aggregation, for determining the final prediction answer (class), and weighted average for aggregating the performance metrics (accuracy, precision, recall) from the clients. Considering that clients might have different dataset sizes, the two methods were chosen to account for the contribution of each client based on its confidence (reliability) and dataset size as the weight, instead of treating all clients equally. The confidence value for each client is calculated from the function predict. This function returns an array of probabilities and associated classes as output. The class with the highest probability in the output array becomes the answer, and its probability becomes the confidence value.

The Bayesian aggregation method by applying parameters α for number of correct predictions (success) and β for number of incorrect predictions (failure), both from the beta distribution, updates and stores cumulative probabilities in a dictionary (class_probabilities) for each class. It then calculates total weight (total_weight) and confidence (reliability) across all clients. Finally, the method normalizes the class_probabilities to obtain the final aggregated probabilities, which are used to determine the final aggregated prediction. From [Ramachandran and Tsokos 2009], If we let H denote our hypothesis and E the evidence (data), the formula can be given as

$$P(H_i|E) = \frac{P(E|H_i)P(H_i)}{\sum_{i=1}^{n} P(E|H_i)P(H_i)} \tag{3}$$

where for each received response i, $P(H_i|E)$ represents the posterior probability, $P(H_i)$ is the prior probability reflecting the initial hypothesis before examining the data and $P(E|H_i)$ is the likelihood, which describes the probability of observing the given data under the hypothesis. The term $\sum_{i=1}^{n} P(E|H_i)P(H_i)$ denotes the evidence, serving as a normalization constant that includes the overall probability of observing the data.

The weighted averaging method uses dataset size from each client as the weight to generate the final weighted performance metrics average. A larger dataset size implies a higher weight for the client's contribution compared to others. For this process, the Equation 1 is applied. See Algorithm 2 for detailed operation.

Algorithm 2 Synchronous FA - 5GLENA Integrated Algorithm

NS-3 5G-LENA simulates the network and collects statistical results of 20 clients
FA Server reads NS-3 network statistical results
for each client i on iteration t in network **do**
  Server selects a subset $S_j^{(t)}$ of clients with $DR_i^{(t)} \geq 90\ \%$
  Server sends a query $q^{(t)}$ to the selected clients in $S_j^{(t)}$
  for each client j in subset $S_j^{(t)}$ **do**
    Client j performs local training of model $\omega_i^{(t)}$ for n epochs
    Client j responds $r_i^{(t)}$ to the query $q^{(t)}$ including performance metrics
  end for
  for each client i not in subset $S_j^{(t)}$ **do**
    Client i sends error response $e_i^{(t)}$ and zero for performance metrics
  end for
  Server aggregates all responses and performance metrics from all clients
end for
Server determines final response

5. Configuration Parameters and Simulation Results

5.1. Simulation Environment

The project is ran on the Ubuntu Linux 64-bit system machine with 8GB of RAM, i5 Quad-core and 2.30GHz CPU frequency. The architecture of the framework consists of integrating two main software blocks, FA and NS-3 with 5G-LENA version 3.3.

5.2. NS-3 and 5G-LENA Configuration

NS-3 and 5G-LENA provide built-in models, protocols, and libraries for simulating different network scenarios. In our experiments, MobilityHelper class is used to position UEs and a gNB (acting as gateway). NodeContainer class manages simulation nodes, including gNB, UEs and other network elements. ThreeGppChannelModel simulates realistic propagation scenarios with path loss, shadowing and fading to accurately model the radio environment. NrHelper and NrPointToPointEpcHelper classes configure 5G New Radio (NR) devices and the core network (EPC), facilitating seamless communication between the gNB and the UEs. Devices are deployed using NetDeviceContainer, InternetStackHelper configures the IP stack for the devices, and IP addresses are assigned through Ipv4InterfaceContainer. The application layer is organized with a UdpServerHelper on a server attached to gNB and UdpClientHelper on remote clients (UEs). For traffic analysis, we use FlowMonitorHelper to collect performance metrics.

5.3. Simulation parameters

The simulation uses various parameters that define the scenario, traffic, and network configurations. As shown in Figure 2, we position a 5G gNB in a fixed location where the FA server is installed and randomly distribute 20 UEs within a radius of 200 meters that must transmit data to the server. Other parameters are presented in Table 2.

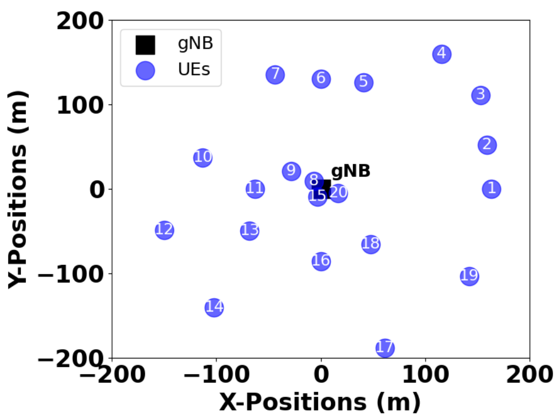

Figure 2. gNB and devices

Parameter	Value
OperationalBand	Single (HT)
UdpPacketSize	1500 bytes
lambdaHT	20000 pps
simTime	1000 ms
udpAppStartTime	400 ms
bandwidthBand	400e6 Hz
numerologyBwp1	3
centralFrequencyBand1	6.8e9 Hz
TxPower	$[0, 30]$ dBm

Table 2. Simulation parameters

The UEs are randomly placed in 2-dimensions around a gNB, which is fixed at the center (origin following the Cartesian coordinates system) with height of 10 meters. This approach guarantees that all devices stay at the same position throughout the simulation and allows better control and fair comparison of the generated results in each configured scenario.

The hypothesis under research is to test whether the optimized power strategy (channel-aware power allocation algorithm) significantly improves both network and FA performance compared to maximum and random power allocation scenarios.The simulation varies transmission power levels between 0 and 30 dBm across three scenarios: (1) uniform maximum power allocation for all UEs, (2) random power allocation, and (3) channel-aware power allocation algorithm. Transmission power values adhere to Anatel resolution [Ana 2021], which regulates wireless communication standards in Brazil. Each scenario undergoes 15 Monte Carlo simulations to ensure statistical reliability. Performance metrics (throughput, latency, packet loss, delivery ratio, and energy consumption) are collected, with confidence intervals and averages calculated and plotted for clear visualization.

5.4. FA Configuration

The FA block consists of a central server that manages system operations, including client selection, query generation, and response aggregation. It connects to the NS-3 core network via a RemoteHost, with the SGW handling data routing and mobility anchoring, while the PGW acts as a gateway to external networks. Bidirectional communication is established using IPv4 (AF_INET) and TCP (SOCK_STREAM) sockets. Two FA algorithms are implemented: one integrates with NS-3 5G-LENA to process network data, while the other includes Bayesian aggregation for final predictions and a weighted averaging method to aggregate client performance results based on reliability.

The FA clients' models utilize a neural network with two fully connected dense layers for MNIST digit classification. The first layer has 512 neurons, and the second has 256, both using ReLU activation to enhance learning efficiency. The final layer consists of 10 neurons with a softmax activation function, converting outputs into a probability distribution for digit classification. The model is optimized with Adam and employs sparse categorical cross-entropy as the loss function, with accuracy, precision, and recall as performance metrics, to ensure effective classification of handwritten digits.

5.5. Simulation Results

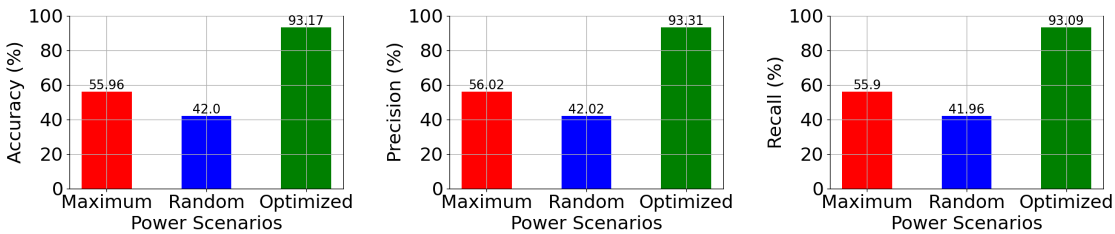

Figure 3. FA performance metrics for power scenarios.

The results of FA simulation on 5G network conditions, given in Figure 3, show differences in performance across three transmission power scenarios. Scenario 1 achieves moderate results in accuracy (55.96 %), precision (56.02 %), and recall (55.90 %). This approach, when using higher power levels, may lead to interference, signal congestion, and inefficient energy use, as it does not account for varying UE conditions. Scenario 2 results in poorer metrics with accuracy: 42.0 %, precision: 42.02 % and recall: 41.96 % due to improper power allocation, which can cause UEs to be either underpowered or overpowered. This may result in reducing network efficiency. In contrast, scenario 3, using a channel-aware power allocation algorithm, gives the best performance with accuracy: 93.17 %, precision: 93.31 %, and recall: 93.09 %. This demonstrates that this adaptive strategy minimizes energy consumption while maintaining reliable communication, including reducing unnecessary waste and potential interference. These findings suggest that channel-aware power allocation, by optimizing power distribution based on network conditions, outperforms both uniform and random power allocation strategies.

The simulation results of FA have revealed that network performance significantly influences the quality of the final aggregated answer. As given in Figure 4, in scenario 1, moderate network performance was observed, including a delivery ratio of 91.69 %, throughput of 224.16 Mbps, and latency of 25.22 ms. However, higher transmission power without considering network conditions might have led to interference and inefficiencies, resulting in 8.31 % packet loss, which likely reduced the number of successful responses sent to the central server. Scenario 2 further worsened performance, with a delivery ratio of 57.51 %, throughput of 140.61 Mbps, and latency of 100.46 ms. The

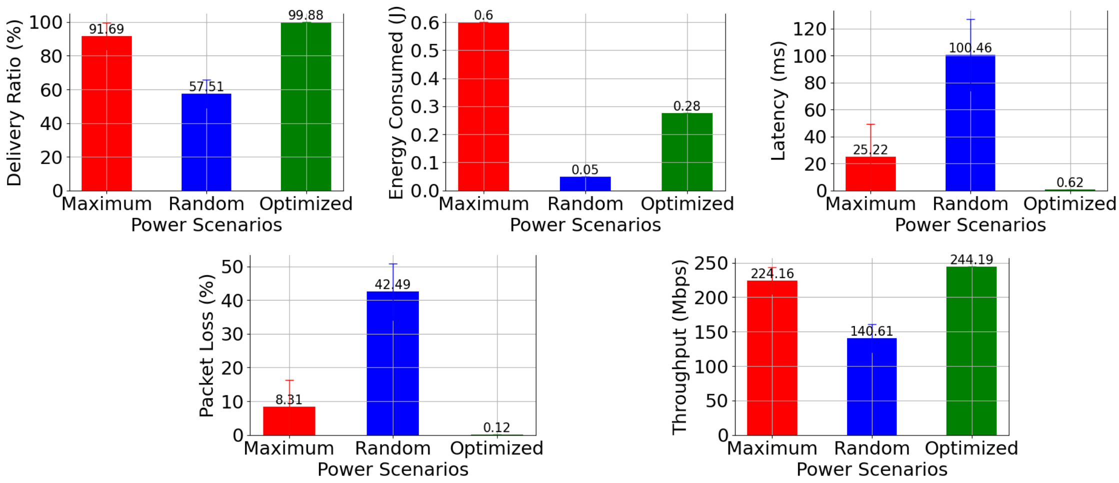

Figure 4. Network performance metrics for power scenarios.

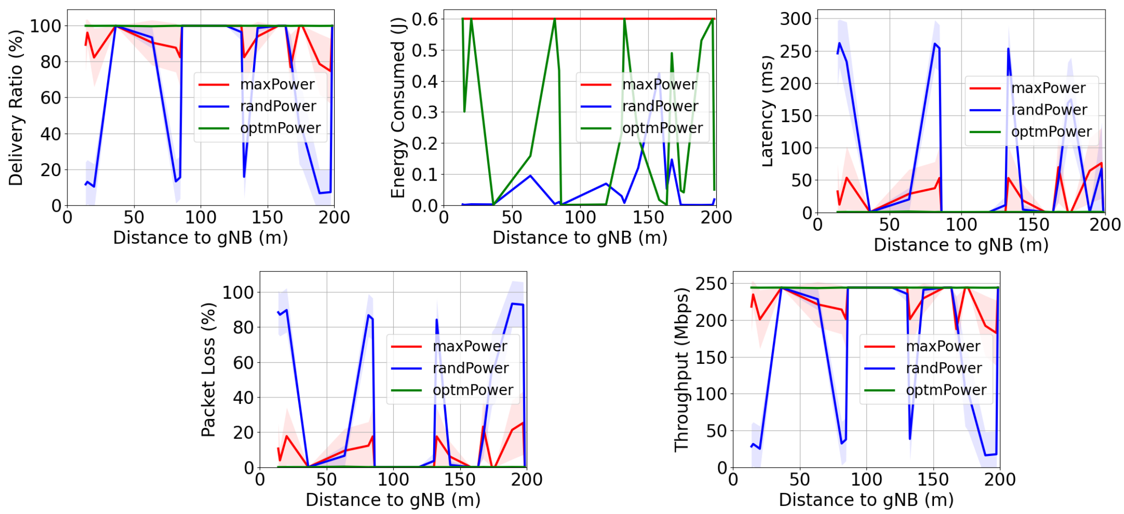

Figure 5. Network performance metrics in terms of distance, for power scenarios.

random approach caused significant packet loss (42.49 %), limiting the number of responses reaching the server, hence, degrading the aggregated results.

In contrast, scenario 3, outperformed the other scenarios with a delivery ratio of 99.88 %, throughput of 244.19 Mbps, and latency of 0.62 ms. This optimized allocation minimized packet loss to 0.12 %, ensuring more successful intermediate responses reached the server, which is important for improving the final FA aggregated result. The algorithm also reduced energy consumption to 0.28 J, which highlights the efficiency of the approach. These results suggest that channel-aware optimization significantly improves network performance, leads to more reliable data transmission and, consequently, more accurate and timely results in FA.

Based on the results shown in Figure 5, within the simulation distance of 200 meters, scenario 3, demonstrates minimal variations in network performance. This shows that the adaptive power allocation helps maintain stability by adjusting transmission power based on real-time conditions such as distance, channel and signal quality. In contrast, scenario 1 and scenario 2 show larger variations. Scenario 1 experiences moderate fluctuations, while scenario 2 exhibits more significant instability. These higher variations might indicate that the lack of dynamic adjustment in these scenarios leads to less consistent performance. This reveals the potential advantage of the channel-aware approach in scenario 3 for achieving more stable results.

Table 3. Statistical significance analysis

Comparison Pairs	Acc.	Prec.	Rec.	DeRa	ThPu	Lat	PaLo	EnCo
Maximum - Optimized	True	True	True	False	False	False	False	True
Maximum - Random	True	True	True	False	False	False	False	True
Optimized - Random	True	True	True	True	True	True	True	False

The statistical significance analysis in Table 3 was performed using ANOVA and Tukey for FA metrics: accuracy (Acc.), precision (Prec.), and Recall (Rec.), since they followed a normal distribution. For network metrics: delivery Ratio (DeRa), throughput (ThPu), latency (Lat.), packet loss (PaLo), and energy consumption (EnCo), the Kruskal-Wallis and Dunn tests were applied as they were not normally distributed. All tests used a significance level of 5 % ($\alpha = 0.05$). "True" in Table 3 indicates the existence of a significant difference between the scenario comparison pairs, while "False" indicates the opposite. The results show that the optimized power allocation significantly improves FA performance compared to both maximum and random power allocation scenarios. However, for network metrics, no significant differences were found between maximum and optimized or maximum and random power allocation scenarios in terms of throughput, latency, delivery ratio, and packet loss, but the optimized scenario significantly reduces energy consumption. This makes the channel-aware power allocation algorithm a better choice for balancing performance and energy efficiency for FA in wireless networks.

6. Conclusion and Future Work

This study demonstrates the significant impact of transmission power strategies on the performance of FA in 5G networks. By comparing three transmission power scenarios: uniform maximum power allocation, random power allocation, and channel-aware power allocation algorithm, it was found that the channel-aware approach consistently outperformed the other two in terms of network stability, reliability, and efficiency. The results suggest that dynamically adjusting transmission power based on network conditions, such as distance, channel quality, and signal conditions, contributes to more reliable communication, and effective transmission of intermediate responses to the central server, which improves the accuracy and reliability of the final aggregated answer in FA system. Future research could integrate additional network metrics, with network data analytics function (NWDAF), refine the power allocation algorithm for varying conditions (interference, congestion, traffic), and explore more energy-efficient strategies to include other state-of-the-art (SOTA) algorithms. Field experiments with diverse user positions and traffic loads could further validate and optimize the proposed approach in real dynamic environments.

References

(2021). Anatel - Resolução nº 747. Agência Nacional de Telecomunicações. Accessed: Jan 2025.

38.901, G. T. (2017). Study on channel model for frequencies from 0.5 to 100 ghz.

Chen, M., Yang, Z., Saad, W., Yin, C., Poor, H. V., and Cui, S. (2021). A Joint Learning and Communications Framework for Federated Learning Over Wireless Networks. *IEEE Transactions on Wireless Communications*, 20(1):269–283.

Elkordy, A. R., Ezzeldin, Y. H., Han, S., Sharma, S., He, C., Mehrotra, S., and Avestimehr, S. (2023). Federated Analytics: A Survey. *APSIPA Transactions on Signal and Information Processing*, 12(1):–.

ETSI (2022). Multi-Access Edge Computing (MEC) Framework and Reference Architecture. *ETSI Standard GS MEC 003*.

Hapanchak, V. S. and Costa, A. D. (2022). Emulation of Multipath Solutions in Heterogeneous Wireless Networks Over Ns-3 Platform. In Jiang, D. and Song, H., editors, *Simulation Tools and Techniques*, volume 424, pages 3–18. Springer International Publishing.

Kairouz, P., McMahan, H. B., Avent, B., Bellet, A., Bennis, M., et al. (2021). *Advances and Open Problems in Federated Learning*. Now Foundations and Trends.

Larrañaga, A., Lucas-Estañ, M. C., Lagén, S., Ali, Z., Martinez, I., and Gozalvez, J. (2023). An open-source implementation and validation of 5G NR configured grant for URLLC in ns-3 5G LENA: A scheduling case study in industry 4.0 scenarios. *Journal of Network and Computer Applications*, 215:103638.

Ramachandran, K. M. and Tsokos, C. P. (2009). *Mathematical Statistics with Applications*. Academic Press.

Shi, S., Hu, C., Wang, D., Zhu, Y., and Han, Z. (2022). Federated Anomaly Analytics for Local Model Poisoning Attack. 40(2):596–610.

Toka, L., Konrád, M., Pelle, I., Sonkoly, B., Szabó, M., Sharma, B., Kumar, S., Annavazzala, M., Deekshitula, S. T., and Franklin, A. A. (2023). 5G on the roads: Latency-optimized Federated Analytics in the Vehicular Edge. *IEEE Access*.

Wang, Z., Ji, H., Zhu, Y., Wang, D., and Han, Z. (2024). A Survey on Federated Analytics: Taxonomy, Enabling Techniques, Applications and Open Issues.

Yue, X., Kontar, R. A., and Gómez, A. M. E. (2024). Federated Data Analytics: A Study on Linear Models. 56(1):16–28.

Zhao, L., Valero, M., Pouriyeh, S., Li, F., Guo, L., and Han, Z. (2024). A Decentralized Communication-Efficient Federated Analytics Framework for Connected Vehicles. pages 1–6.

Zhao, L., Valero, M., Pouriyeh, S., Li, L., and Sheng, Q. Z. (2022). Communication-Efficient Semihierarchical Federated Analytics in IoT Networks. 9(14):12614–12627.

Zárate Ceballos, H., Parra Amaris, J. E., Jiménez Jiménez, H., Romero Rincón, D. A., Agudelo Rojas, O., and Ortiz Triviño, J. E. (2021). *Wireless Network Simulation: A Guide Using Ad Hoc Networks and the NS-3 Simulator*. Apress.