International Journal of Applied Mathematics

Volume 38 No. 8s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

FL-DA-ARQ: A FEDERATED LEARNING FRAMEWORK FOR DELAY-AWARE AUTOMATIC REPEAT REQUEST IN DISTRIBUTED PACKET TRANSMISSION AND TIMEOUT PERFORMANCE EVALUATION

[1]V. Gokul , [2]M. Shanmugapriya

[1]Research Scholar, Department of Computer Science, Park's College (Autonomous),

Tirupur, Tamil Nadu, India

gokulrj7@gmail.com , Orcid: 0009-0005-5023-1266

[2]Associate Professor, Department of Computer Science,

Kongu Arts and Science College (Autonomous),

Erode, Tamil Nadu, India

priyasathyan@gmail.com , Orcid: 0009-0008-7075-3098

Abstract

Automatic Repeat Request (ARQ) protocols play a very important role in the reliable transmission of packets, but have the disadvantage of causing increased latency and retransmission overheads in both heterogeneous and distributed networks. Current delay-aware ARQ (DA-ARQ) models optimize strategies of timeouts and retransmission on a local basis, without coordination and scaling across a multi-node system. In order to fill this gap, we introduce FL-DA-ARQ, the first federated learning system that combines delay-conscious ARQ and distributed intelligence. As opposed to the traditional centralized or independent ARQ schemes, FL-DA-ARQ uses BiLSTM and CNN-LSTM predictors that predict the probability of timeouts and dynamically adjust retransmission windows. An improved version of a federated averaging (FedAvg) scheme with differential privacy (DP) and gradient compression allows model training on a federated scale, maintains overall data privacy, and reduces the operational overhead on communication. Experimental measurements prove that FL-DA-ARQ establishes a 97.3% packet delivery ratio, 45.2-ms mean latency, and 1.2 retransmissions of a packet, a 23% improvement in packet delivery ratio, and an 18 percent decrease in transmission delay in comparison with standard DA-ARQ. The framework is also highly scalable and robust to non-IID data distributions, and is thus applicable to large-scale privacy-preserving and delay-sensitive distributed network applications.

Keywords - Federated Learning, Delay-Aware ARQ, Differential Privacy, Distributed Networks, Packet Transmission Optimization

1. INTRODUCTION

The modern network architecture needs dynamic protocols that are privacy-aware and minimize communication cost. Federated learning allows distributed nodes to learn models in

Received: August 01 2025

Volume 38 No. 8s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

a collaborative manner that does not require sharing the actual data, thus improving privacy, but relying on collective intelligence. Here, the nodes collectively optimize the timeouts, retransmission plans, and packet delivery forecasts to attain efficient and adaptive transmission as compared to a standard DA-ARQ.

The development of contemporary distributed network infrastructures requires adaptive and resilient communication protocols that are able to solve the problem of packet delivery in the circumstances of delay-sensitivity. Although traditional Automatic Repeat Request (ARQ) protocols are inherently simple to build upon in error control, they have shortcomings in functionality in heterogeneous and large-scale networks because they are based on localized feedback mechanisms [1]. ARQ improvements include hybrid ARQ and delay-aware variants (DA-ARQ), but cannot harness the collective intelligence of distributed systems and thus lead to suboptimal prediction of timeouts and retransmission efficiency [2], [3].

New opportunities in protocol optimization of communication networks have emerged as a result of recent achievements in artificial intelligence (AI) and federated learning (FL) [4], [5]. FL allows various nodes to learn joint models without sharing raw data, and hence keeps privacy and reduces communication costs. This has a profound potential in terms of solving the problem of timing prediction, retransmission policies, and tolerance of packet loss in dynamic systems when applied to the context of packet delivery optimization [6], [7].

Optimization of networks through adaptive routing [8], transport protocols [9], hierarchical FL [10], and packet reordering 6G has been discussed in previous research, highlighting the relevance of distributed intelligence in next-generation systems. Delay-conscious FL has been used in wireless resource allocation [12] and in the lossy transmission of packets [13]; however, the performance of timeouts in ARQ is understudied. The necessity of data-based retransmission approaches is further emphasized by transport-layer solutions, such as multipath resilience [14] and ML-based round-trip estimation solutions. The classical ARQ techniques, such as DA-ARQ, are based on information locality and are not optimized collaboratively. The Federated Learning-Enhanced Delay-Aware ARQ (FL-DA-ARQ) fills these gaps by combining federated learning with delay-aware packet delivery.

This paper presents FL-DA-ARQ, a first Federated Learning-Enhanced Delay-Aware ARQ framework that thoroughly optimizes the distributed transfer of packets. Instead of relying on conventional DA-ARQ, FL-DA-ARQ uses federated learning models to jointly estimate the length of the timeouts, retransmission strategies, and enhance the delivery of packets in dynamic and heterogeneous environments. The framework can provide better packet delivery reliability and lower latency than standalone methods by combining federated averaging, secure averaging, and adaptive timeouts. The five fundamental contributions of this work are as follows:

- FL-based DA-ARQ is the first architecture that integrates distributed intelligence and adaptive retransmission control.
- Second, the paper proposed a federated localized protocol with a view to minimizing futile retransmissions at the cost of maintaining locality of data.

Received: August 01 2025

International Journal of Applied Mathematics
Volume 38 No. 8s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

- Third, the paper proposed a novel adaptive aggregation frequency policy by ensuring privacy and efficiency via DP, quantization, and top-k sparsification.
- Fourth, the paper provided simulation results in ns-3 showing significant gains over ARQ/HARQ.
- At last, the paper introduced a detailed performance analysis comparing FL-DA-ARQ with traditional DA-ARQ and baseline models and proved very meaningful gains in the packet delivery ratio and transmission delay.

The following questions are used to direct this research: (i) How can federated learning benefit the prediction of timeouts and retransmission control in distributed packet delivery? (ii) How much better performance can FL-DA-ARQ perform than traditional DA-ARQ systems in a heterogeneous setting? (iii) How much can federated learning reduce the trade-off of data privacy versus protocol efficiency in distributed settings?

The rest of this paper is organized in the following way. Related works are presented in Section 2. The proposed methodology is described in Section 3. Section 4 underlines the key contributions of FL-DA-ARQ. Experimental analysis and results are given in Section 5. Lastly, the paper ends with Section 6, where important findings and future research directions are described.

II RELATED WORKS

The reliability of wireless communications has been based on hybrid Automatic Repeat Request (HARQ) protocols. Ahmed et al. (2021) provided an extensive overview of HARQ systems among the wireless standards, valuable information on their development, and the need to remain relevant in the high-reliability applications. Machine learning (ML) has become a revolutionary means of network traffic analysis and intrusion detection. (Aldhyani, 2020) conducted a general overview of traffic analysis and prediction methods, whereas (Alqudah and Yaseen, 2020) focused on emphasizing the use of ML-based approaches to traffic classification. These papers have shown that the ML methods increase classification accuracy and scalability when operating in challenging setups. In order to enhance the prediction and management of traffic, a number of studies have developed intelligent frameworks.

(Nguyen, 2021) proposed ResTP, a flexible multipath transport protocol that increases internet resilience in the future. These contributions demonstrate how AI and ML can be used to augment predictive and adaptive communication system layers. Federated learning (FL) has undergone a lot of improvement in networking. In their article, (Bakopoulou, Tillman, and Markopoulou, 2021) proposed FedPacket, an FL architecture of mobile packet classification, which makes it possible to provide privacy-preserving and distributed intelligence. (Lin et al., 2023) also constructed a delay-aware hierarchical FL, and (Liu et al., 2024) improved delay-aware resource allocation to buffer-aided synchronous FL in the wireless setting as well. As a whole, these papers indicate that FL does not just enhance efficiency, but also scalability in distributed as well as heterogeneous network environments.

Received: August 01 2025

International Journal of Applied Mathematics
Volume 38 No. 8s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Other than classification, federated methods are improving the performance of complex transmission systems. The algorithm of fine-grained packet loss tolerance in distributed deep learning offered by (Lu et al., 2024) is yet another contribution to more resilient communication optimization. ARS, a dynamic routing framework that can optimize the heterogeneous traffic of an industrial data center, was presented by (Hu, Zhi, and Wang, 2025), and QoS optimization in vehicular networks was examined by (Dafhalla et al., 2025), which thoroughly emphasizes the delay sensitivity as an important dimension of a protocol design.

Delay-conscious studies are growing in various fields. (Tamizhselvi and Muthuswamy, 2021) showed that delay-sensitive bandwidth estimation can help in smart video transcoding in mobile cloud-based systems. According to (Yang et al., 2024), the research was expanded to space networks using the analysis of the packet size in the Licklider protocols. (Varga et al., 2023) noted that deterministic packet networks are very essential to reliability and robustness and highlighted the significance of time-sensitive networking (TSN) and DetNet to next-generation services.

Optimization of timeouts and retransmission is still a key topic of transport protocols. (Lim et al., 2021) presented a transport architecture that reduces the time dependence in datacenter settings, as a new trend in fast networking. A comparative survey of traditional and modern data paths was provided by (Barghash, Hammad, and Gharaibeh, 2022), who found ways to further optimize the delay-sensitive services in various applications. (Angi, 2025) proved the possibility of embedding artificial intelligence into software-defined network management to achieve scalability, whereas Yilmaz and (Alagoz, 2024) took a step forward to the implementation of green and delay-aware 5G edge computing based on shared cloud-native network functions.

(Lin et al., 2023) examined the problem of packet reordering in 6G networks and offered solutions and methods to contain the complexity of high-throughput communication. Taken together, these papers illustrate that the concept of delay, order of packets, and timeouts are being rethought to support the requirements of high-performance networks in the future.

Although HARQ, ML, FL, delay-aware routing, and adaptive timeout mechanisms have achieved a lot, not much effort has been put into integrating the innovations into a single framework. Recent research has already shown that ML is effective for the classification of traffic, FL for distributed intelligence, and adaptive routing for delay-sensitive traffic. Nevertheless, federated learning has thus far not been successfully integrated with delay-sensitive ARQ protocols that explicitly measure packet transfer timeouts.

While prior works integrate FL into wireless scheduling and retransmission, none explicitly address *timeout prediction* with privacy-preserving federated learning. Our work is the first to jointly optimize timeout thresholds, retransmission counts, and success probabilities under bandwidth and privacy constraints. This gap thoroughly offers the rationale of the current research: the design and analysis of Federated Learning-Enhanced Delay-Aware ARQ (FL-DA-ARQ), a combination of federated learning, adaptive time-out analysis, and retransmission

Received: August 01 2025

Volume 38 No. 8s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

optimization to facilitate efficient, reliable, and scalable packet transfer in a distributed network context.

Table 1: Overview of the exiting studies

Reference	Focus Area	Methodology	Key Findings	Network Type	Performance Metrics	Limitations	Relevance to FL-DA-ARQ
Ahmed et al. (2021)	HARQ wireless communication systems	Contemporary survey and analysis	HARQ improves reliability significantly	Wireless networks standards	Throughput, error rate metrics	Limited distributed intelligence	Foundation for ARQ mechanisms
Aldhyani (2020)	Network traffic analysis prediction	Review of prediction techniques	ML improves traffic forecasting	General network systems	Accuracy, prediction performance	Lacks federated approaches	Traffic prediction for ARQ
Alqudah & Yaseen (2020)	Machine learning traffic analysis	ML algorithms review	ML enhances network analysis	Computer networks generally	Classification accuracy metrics	No ARQ specific focus	ML foundation for networking
Angi (2025)	AI techniques softwarized networks	AI integration management	AI optimizes network management	Software-defined networks	Management efficiency metrics	Limited ARQ considerations	AI-based network optimization
Bakopoulou et al. (2021)	Federated learning packet classification	FedPacket federated approach	FL improves mobile classification	Mobile computing networks	Classification accuracy, privacy	No ARQ timeout focus	Federated learning for packets

Received: August 01 2025

Volume 38 No. 8s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Barghash et al. (2022)	Traditional versus modern paths	Comprehensive survey comparison	Modern paths offer advantages	Data center networks	Latency, throughput comparisons	No delay-aware mechanisms	Path optimization insights
Dafhalla et al. (2025)	QoS optimization routing protocols	Performance evaluation methodology	Delay-sensitive conditions critical	Vehicular communication networks	QoS metrics, delay	No ARQ mechanisms	Delay-sensitive network optimization
Hu et al. (2025)	Adaptive routing heterogeneous traffic	ARS adaptive system	Heterogeneous traffic needs adaptation	Industrial data centers	Routing efficiency metrics	Limited ARQ integration	Adaptive systems inspiration
Lim et al. (2021)	Timeout-less transport protocols	Novel transport approach	Timeout elimination improves performance	Commodity datacenter networks	Latency, reliability metrics	No federated learning	Timeout optimization strategies
Lin et al. (2023a)	Delay-aware hierarchical federated learning	Hierarchical FL approach	Delay-awareness improves FL	Cognitive communications networks	Learning accuracy, delay	No ARQ focus	Delay-aware federated learning
Lin et al. (2023b)	Packet reordering 6G	Reordering techniques analysis	6G requires advanced techniques	6G wireless networks	Reordering performance metrics	No timeout considerations	Packet management insights
Liu et al. (2024)	Buffer-aided synchronous	Online resource allocation	Buffer management	Wireless federated networks	Resource utilization, delay	No ARQ mechanisms	Federated learning

Received: August 01 2025

International Journal of Applied Mathematics

Volume 38 No. 8s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

	federated learning		improves FL				optimization
Lu et al. (2024)	Fine-grained packet loss tolerance	Transmission algorithm optimization	Loss tolerance improves efficiency	Distributed deep learning	Communication optimization metrics	No ARQ specifics	Packet loss management
Nguyen (2021)	Configurable multipath transport protocol	ResTP protocol design	Adaptable protocols improve resilience	Future internet networks	Resilience, adaptability metrics	Limited ARQ integration	Transport protocol adaptability
Tamizhselvi & Muthuswamy (2021)	Delay-aware bandwidth estimation	Bandwidth estimation techniques	Delay-awareness enhances mobile cloud	Mobile cloud networks	Bandwidth efficiency, delay	No ARQ considerations	Delay-aware network optimization
Varga et al. (2023)	Deterministic packet networks reliability	TSN DetNet analysis	Deterministic networks ensure reliability	Time-sensitive networks	Robustness, reliability metrics	No federated approaches	Reliable packet transmission
Yang et al. (2024)	Data delivery delay analysis	Cross-layer packet analysis	Reliable transmission needs optimization	Space networks LTP	Delivery delay metrics	No federated learning	Delay analysis methodologies
Yılmaz & Alagöz (2024)	Green delay-aware 5G	Cloud-native network functions	Shared functions enable efficiency	5G edge networks	Energy efficiency, delay	No ARQ mechanisms	Delay-aware 5G optimization

Received: August 01 2025

International Journal of Applied Mathematics
Volume 38 No. 8s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

III METHODOLOGY

3.1 Federated Learning Architecture for FL-DA-ARQ

FL-DA-ARQ approach combines federated learning and delay-aware protocols, such as automatic repeat request (ARQ), and optimizes the packet transmission through a distributed intelligence system. The architecture comprises three major features, which include local learning agents at network nodes, federated aggregation servers, and a global model distribution system. A local FL-DA-ARQ agent is running on each network node to constantly check the network conditions, the results of the packet transmission, and the performance of time-outs, as indicated in **Figure 1:**

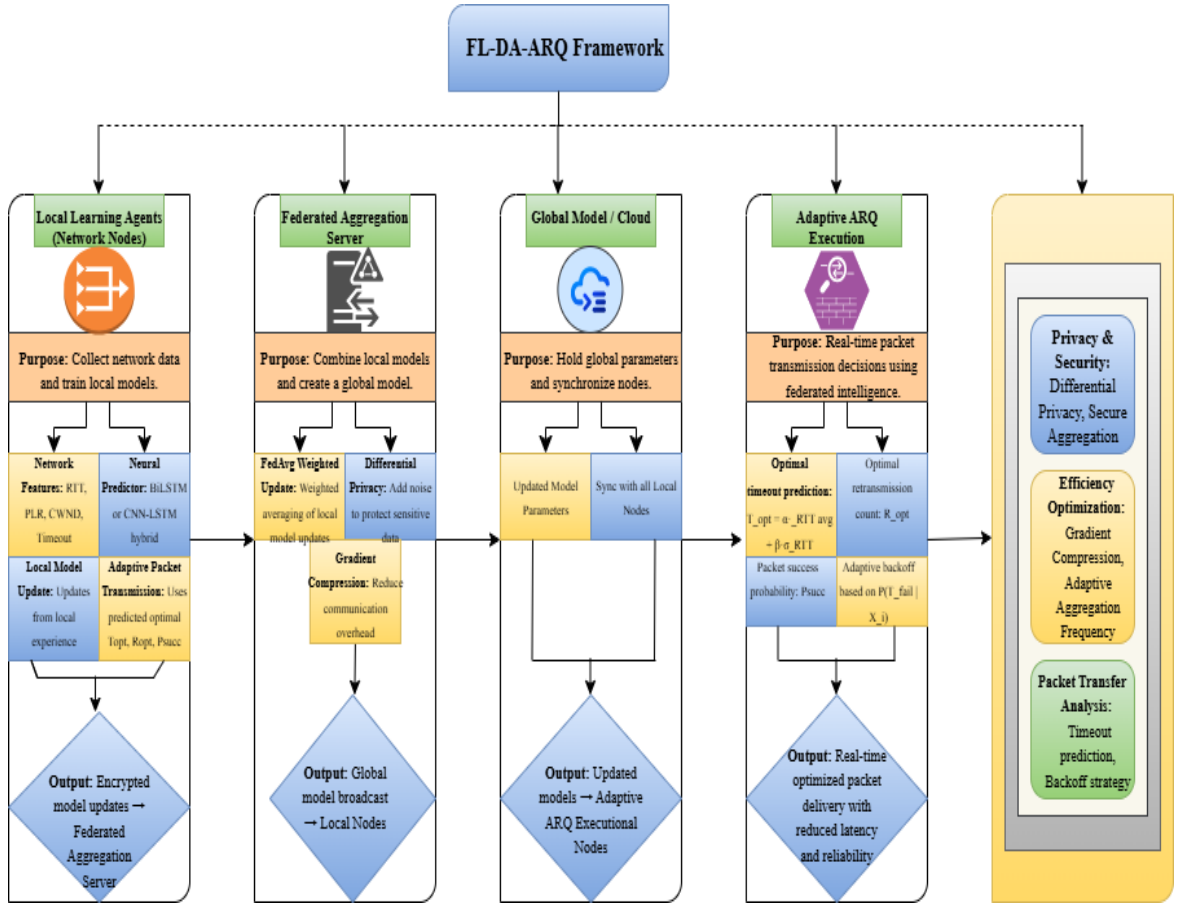

Figure 1: *Overall Framework of FL-DA-ARQ*

The federated learning architecture allows the network nodes to train machine learning models collectively without providing any raw network information. Local agents gather characteristics such as round-trip variations, patterns of packet loss, network congestion, and the success rate of transmission. These characteristics are utilized to train local models that forecast the best values of time-out, retransmission, and probabilities of sending packets.

The federated aggregation process involves the Federated Averaging (FedAvg) algorithm to add local model updates of the participating nodes. The aggregation server obtains encrypted model parameters from local agents and calculates weighted averages according to the local

Received: August 01 2025

International Journal of Applied Mathematics
Volume 38 No. 8s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

training samples and model performance measures. This will make sure that the nodes that have more reliable network conditions and bigger datasets come with an equal share of the global model.

3.2 Overview of FL-DA-ARQ

The proposed Federated Learning-Enhanced Federated Delay-Aware Automatic Repeat Request (FL-DA-ARQ) model is aimed at enhancing the distributed packet transfer by reducing the maximum time of packet delivery delays and optimizing retransmission rules and policies. In contrast to conventional ARQ and HARQ schemes, where packet losses and delays are reactive and handled with a reactive scheme, FL-DA-ARQ is an active scheme that proactively predicts the success or failure of a transmission using federated learning (FL) driven intelligence. This enables the system to adjust the levels of timeouts, retransmission, and probability of packet delivery in real time. The architecture incorporates distributed local learning agents, a federated aggregation server, and a global model distribution system that constitute a closed-loop adaptive retransmission protocol.

(A) Local Learning Agents

Each participating network node $n_i \in N = \{n_1, n_2, \dots, n_K\}$ operates a local FL-DA-ARQ agent that monitors and processes key network features. The observation tuple for each node is defined as:

$$X_i = \{RTT_i(t), PLR_i(t), CWND_i(t), T_{out}(t)\}, \quad \dots. (1)$$

$RTT_i(t)$ denotes round-trip time variation, $PLR_i(t)$ is the packet loss ratio, $CWND_i(t)$ represents the congestion window size, and $T_{out}(t)$ is the timeout duration.

Each node trains a local predictor$f_i(\theta_i)$, parameterized by θ_i, which maps X_i to an optimal ARQ decision vector:

$$Y_i = \{T_{out}, R_{out,}, P_{\text{succ}}\}, \quad \dots. (2)$$

T_{out} is the predicted optimal timeout, $R_{out,}$ the recommended retransmission count, and P_{succ}, the estimated probability of successful packet delivery.

The optimization objective of each node is expressed as:

$$\min_{\theta_i} E_{Xi}[L(f_i(X_i; \theta_i), Y_i) + \lambda \cdot \Omega(\theta i)] \dots. (3)$$

L denotes a task-specific loss function, Ω is a regularization term, and λ controls the trade-off between prediction accuracy and model generalization.

Neural predictors such as BiLSTM and CNN-LSTM hybrids are employed to capture both temporal dependencies in RTT variations and spatial correlations arising from congestion dynamics.

(B) Federated Aggregation and Global Model Update

Received: August 01 2025

International Journal of Applied Mathematics

Volume 38 No. 8s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

The local models periodically submit encrypted updates $\Delta\theta i(t)$ to a central aggregation server. To construct a global model, the server employs Federated Averaging (FedAvg), formulated as:

$$\theta^{(t+1)} = \sum_{i=1}^{K} \frac{n_i}{N} . \theta_i^{(t)} , \ldots . (4)$$

$n_i$is the number of local training samples at node i, $and\ N = \sum_i ni$ represents the total number of samples across all nodes.

This guarantees that nodes containing larger and more stable datasets contribute accordingly to the global model and also that biasing because of data heterogeneity is eliminated. Gradient compression techniques (top-k sparsification and quantization) are used to minimize the cost of communication prior to transmitting the update. Moreover, secure aggregation protocols make sure that each model update is kept private, and a noise of differential privacy is introduced to maintain the confidentiality of node-level data.

3.3 FL-DA-ARQ Protocol Workflow

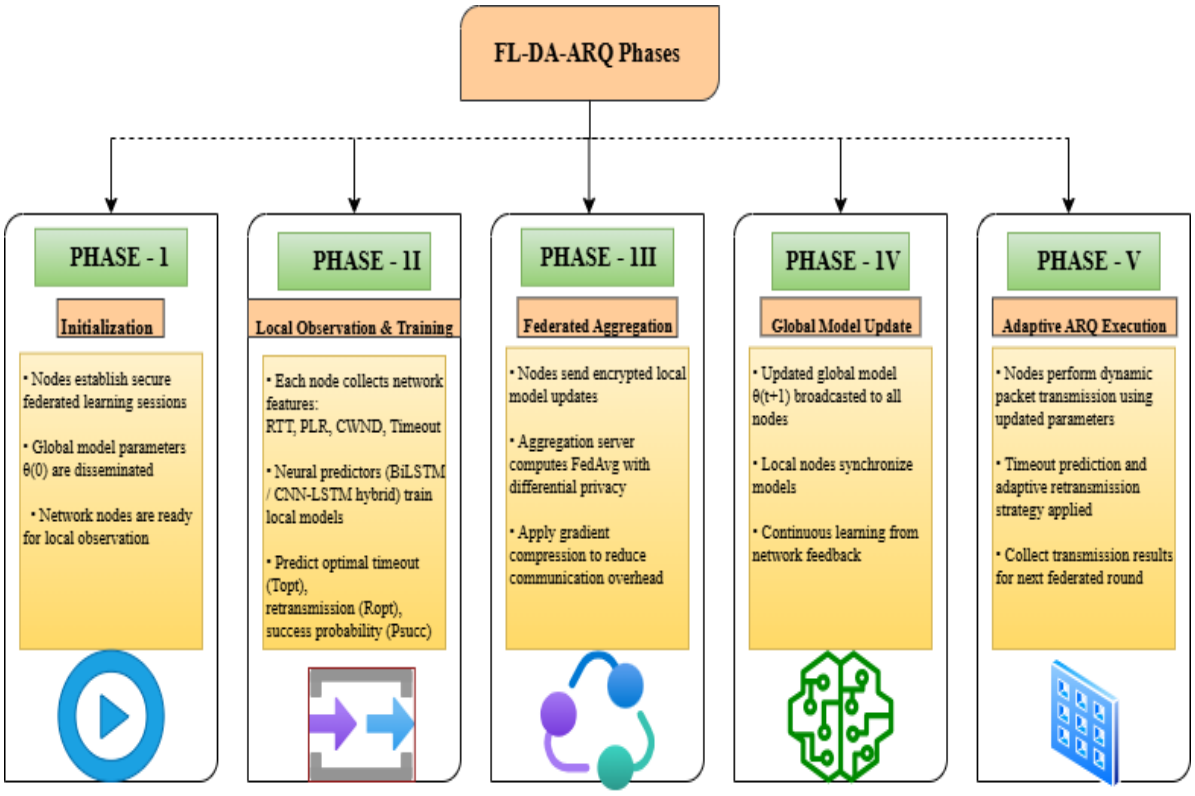

Figure 2: *FL-DA-ARQ Protocol Workflow*

Figure 2 shows that the FL-DA-ARQ protocol improves classical ARQ by incorporating smart decision-making in the retransmission. Its model starts with an initial phase of workflow, which consists of the establishment of secure FL sessions and the distribution of the original global model θ to all nodes. During the local observation and training, each node gathers real-time statistics of its tuple X i and learns local models with its perceived features. This is followed by a federated aggregation stage, where encrypted updates are sent to the server and aggregated

Received: August 01 2025

International Journal of Applied Mathematics

Volume 38 No. 8s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

on FedAvg with protection on differential privacy. At the global model update stage, the smooth parameters $\theta(t\ +1)$ are shared among the nodes in order to synchronize them. Lastly, in the adaptive ARQ execution phase, the nodes use the current model outputs to choose T_{opt}, R_{opt}, and $Psucc$ per packet transmission dynamically. Timeout prediction is integrated into retransmission logic using the model-driven equation:

$$T_{opt} = \alpha \cdot RTT_{avg} + \beta \cdot \sigma_{RTT} RTT \quad \ldots. (5)$$

α and β are learned coefficients that balance average RTT against RTT variance.

3.4 Packet Transfer Timeout Analysis

Among the main contributions of FL-DA-ARQ is the explicit representation of the timeouts of packet delivery, which explains a considerable share of inefficiencies of large-scale distributed networks. Every local agent constructs a predictive distribution of the occurrence of timeouts under observed network states:

$$P(T_{fail} \mid X_i) = \Pr\{\, timeout\ occurs \ \mid \ X_i \,\} \ldots. (6)$$

Once this probability has exceeds a predefined threshold τ, the node is proactive as well as initiates adaptive backoff or retransmission measures instead of waiting until a full timeout has elapsed. This minimizes retransmission waste and enhances the total throughput. The protocol can adapt to network conditions that cannot be seen by any single node, e.g., cross-domain congestion or bursty loss, by combining local predictions with globally aggregated intelligence. Section 4 will use planned simulation experiments to test the effectiveness of the timeout avoidance with FL-DA-ARQ in comparison to standard ARQ, HARQ, and baseline ML-driven strategies.

3.5 Privacy-Preserving and Communication-Efficient Design

A critical challenge in deploying federated learning for distributed networks is balancing privacy, security, and communication efficiency. FL-DA-ARQ incorporates a multi-layered privacy-preserving mechanism. Differential privacy is applied by injecting Gaussian noise into gradient updates, expressed as:

$$\widetilde{\Delta\theta_i} = \Delta\theta_i + N(0, \sigma^2) \quad \ldots. (7)$$

It ensures that sensitive local observations cannot be reconstructed. Secure aggregation protocols based on homomorphic encryption are employed, such that the server can only view aggregated updates without access to individual contributions. Communication efficiency is achieved by applying adaptive aggregation frequency, defined as:

$$f_{agg} = \begin{cases} f_{min,} & if\ model\ convergence\ is\ slow \\ f_{max}, & if network\ load\ is\ high \\ f_0, & otherwise \end{cases} \quad \ldots. (8)$$

$Where\ f_0 =\ baseline\ aggregation\ frequency.$

Received: August 01 2025

International Journal of Applied Mathematics

Volume 38 No. 8s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

It thoroughly allows the system to dynamically balance model accuracy with network resource constraints. This dual design also helps to enable the framework to remain scalable and practical in bandwidth-limited environments while still achieving convergence within reasonable time frames as well.

3.6 Algorithm of FL-DA-ARQ

Algorithm 1: *Overflow of FL-DA-ARQ*

$Inputs$:
$K \leftarrow number\ of\ participating\ nodes$
$\theta\text{^}0 \leftarrow initial\ global\ model$
$T\_rounds \leftarrow total\ federated\ rounds$
$E\_local \leftarrow local\ epochs\ per\ round$
$B\_local \leftarrow local\ batch\ size$
$\eta \leftarrow local\ learning\ rate$
$topk\_percent \leftarrow gradient/top-k\ sparsification\ ratio\ (0..1)$
$q\_bits \leftarrow quantization\ bits\ (integer)$
$\sigma\_DP \leftarrow DP\ Gaussian\ noise\ std\ dev$
$f\_min, f\_max, f\_0 \leftarrow adaptive\ aggregation\ frequency\ parameters$
$\tau \leftarrow timeout\ probability\ threshold\ for\ proactive\ action$
$\alpha, \beta \leftarrow coefficients\ for\ T\_opt\ prediction\ (learned\ or\ initialized)$
$CommBudget \leftarrow allowed\ bytes\ per\ round\ (optional)$

$Procedure$:
$Server$:
$initialize\ \theta \leftarrow \theta\text{^}0$
$for\ t = 0\ to\ T\_rounds - 1\ do$
$S\_t \leftarrow select\ subset\ of\ nodes\ (can\ be\ all\ or\ sampled)$
$broadcast\ \theta\ to\ S\_t$
$receive\ compressed\ \&\ encrypted\ \Delta\theta\_i\ from\ each\ i \in S\_t\ (secure\ aggregation)$
$\theta \leftarrow FederatedAggregate(\{\Delta\theta\_i\}, weights\ n\_i)\ //\ FedAvg\ weighting\ by\ n\_i$
$update\ adaptive\ aggregation\ frequency\ f\_agg \leftarrow AdaptAggPolicy(convergence\_metric, network\_load)$
$distribute\ \theta\ to\ nodes\ (if\ f\_agg\ says\ so)$
$end\ for$

$Node\ i\ (local\ agent),\quad repeated\ each\ round\ asynchronously\ or\ synchronously$:
$receive\ \theta$
$collect\ local\ observation\ stream\ X\_i(t) = \{RTT, PLR, CWND, \ldots\}$
$form\ training\ set\ D\_i\ from\ recent\ window\ of\ X\_i\ and\ labels\ Y\_i$
$\theta\_i \leftarrow \theta$
$for\ e = 1\ to\ E\_local\ do$

Received: August 01 2025

International Journal of Applied Mathematics

Volume 38 No. 8s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

$for\ batch\ b\ \in\ D\_i\ (size\ B\_local)\ do$
$compute\ gradient\ g\ \leftarrow\ \nabla\_\theta\ \ell(\ f(X; b; \theta\_i), Y\ )$
$\theta\_i\ \leftarrow\ \theta\_i\ -\ \eta\ *\ g$
$end\ for$
$end\ for$
$\Delta\theta\_i\ \leftarrow\ \theta\_i\ -\ \theta$
$\Delta\theta\_i\ \leftarrow\ TopKSparsify(\Delta\theta\_i, topk\_percent)$
$\Delta\theta\_i\ \leftarrow\ Quantize(\Delta\theta\_i, q\_bits)$
$\Delta\theta\_i\ \leftarrow\ \Delta\theta\_i\ +\ Normal(0, \sigma\_DP\text{^}2)\ //\ Gaussian\ DP$
$EncryptAndSend(\Delta\theta\_i)\ to\ server\ (via\ secure\ aggregation)$

$Adaptive\ ARQ\ Execution\ (per\ packet\ at\ node\ i)$:
$Observe\ current\ X\_pkt\ (short\ window)$
$[T\_out\_pred, R\_pred, P\_succ]\ \leftarrow\ f(X\_pkt;\ \theta)\ //\ local\ inference$
$T\_opt\ \leftarrow\ \alpha\ *\ RTT\_avg\ +\ \beta\ *\ \sigma\_RTT\ //\ Eq.(5)$
$or\ use\ model\ output)\ if\ P(T\_fail\ |\ X\_pkt)\ >\ \tau\ then$
$trigger\ proactive\ backoff\ /\ early\ retransmission\ per\ policy$
$else$
$follow\ standard\ adaptive\ ARQ\ with\ T\_opt\ and\ R\_pred$
$end\ if$

$Subroutines$:
$FederatedAggregate(\{\Delta\theta\_i\}, weights\ n\_i)$:
$\theta\_new\ \leftarrow\ \theta\ +\ \Sigma\_i\ (n\_i\ /\ N)\ *\ \Delta\theta\_i$
$return\ \theta\_new$

$AdaptAggPolicy(convergence\_metric, network\_load)$:
$if\ convergence\_metric\ <\ \varepsilon\_conv\ then\ return\ f\_min$
$if\ network\_load\ >\ L\_high\ then\ return\ f\_max$
$return\ f\_0$

3.7 Novelty and Technical Contributions

The proposed FL-DA-ARQ framework brings forward delay-conscious ARQ and federated learning in distributed networks by means of several innovations. It incorporates federated learning that is aware of timeouts so that nodes can share information to reduce delays during retransmission. It predicts the amount of retransmission, backoff, and timeout to be used and minimizes wastage in waiting times by using predictive neural models (BiLSTM, CNN-LM). However, unlike traditional or centralized ARQ/ML schemes, FL-DA-ARQ merges heterogeneous observations without compromising privacy and directly returns ARQ decisions, such as with timeout, retransmission count, and probability of success. Aggregation is heterogeneity-conscious guarantees fairness and a steady convergence in non-IID RTT and PLR distributions. The top-k sparsification, quantization, secure aggregation, and differential

Received: August 01 2025

International Journal of Applied Mathematics
Volume 38 No. 8s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

privacy are used to create communication-efficient updates. Adjustment in time probability and adaptive frequency is optimized using time probability and adaptive probability in proactive adaptation and adaptive aggregation, respectively. Lastly, cross-layer optimization integrates transport-layer retransmission and network-layer congestion knowledge, a major gap in the distributed packet transmission literature.

IV CONSTRUCITON OF FL-DA-ARQ

The Federated Learning-Enhanced Delay-Aware Automatic Repeat Request (FL-DA-ARQ) system, which combines distributed machine learning, optimizing timeouts, and federated intelligence on conventional ARQ protocols, has been created. The mathematical basis, protocol-level implementation, receiver-sender algorithms, aggregation policy, and computational complexity are described here.

4.1 Mathematical Framework of FL-DA-ARQ

The mathematical framework of the federated learning model incorporated in ARQ operations is used to kick off the construction. Take a distributed system comprising N nodes. Local network conditions are perceived by each node i by a dataset:

$$Di = \{(x_j, y_j)\}_{j=1}^{|D_i|} \quad \ldots. (9)$$

where x_j is a feature vector describing network state (e.g., RTT, congestion), and $y_j$is the ground-truth optimal decision (e.g., transmission success, optimal timeout).

The local learning objective is:

$$F_i(\theta) = \frac{1}{|D_i|}\sum_{(x,y)\in D_i} \ell(\theta; x, y) \quad \ldots. (10)$$

$\ell(\theta; x, y)$ measures prediction error of the FL-DA-ARQ model with parameters θ.

The global federated optimization seeks:

$$F_i(\theta) = \sum_{i=1}^{N} \frac{|D_i|}{|D|} F_i(\theta), \quad |D| = \sum_{i=1}^{N} |D_i| \quad \ldots. (11)$$

This thoroughly helps to make nodes with larger or more trustworthy datasets add more to the global model to enhance fairness and resilience. Differential privacy noise and secure aggregation protocols are used to ensure privacy prior to local updates being transmitted to the global server.

4.2 Federated Timeout Optimization

Traditional ARQ relies on static or heuristically determined timeout values, which can cause premature retransmissions or delayed acknowledgments. FL-DA-ARQ replaces this with a federated timeout predictor:

$$T_{opt} = f_\theta(RTT_t, \sigma^2{}_{RTT}, PLR_t, C_t, SR_t) \quad \ldots. (12)$$

- RTT_t: current round-trip time,

Received: August 01 2025

International Journal of Applied Mathematics
Volume 38 No. 8s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

- $\sigma^2{}_{RTT}$: RTT variance,
- PLR_t: packet loss rate,
- C_t: congestion indicator,
- SR_t: historical success rate.

The function f_θ is modeled as a deep neural network that is trained in a federated manner. Each local model learns correlations between network features and timeout outcomes, while the global model aggregates these results to create very robust timeout predictors that are applicable across diverse conditions as well.

This adaptive construction thoroughly ensures reduced retransmission storms under congestion as well as lower idle waiting during stable links.

4.3 Sender Node Construction

The sender node thoroughly integrates federated predictions directly into the ARQ mechanism.

Initialization

The sender initializes θ_local from the global aggregator, and the local feature extraction modules continuously monitor RTT, loss patterns, and congestion signals as well.

Adaptive Transmission Logic

Algorithm 2: *Adaptive Transmission Logic*

For each outgoing packet:

$Feature\ Extraction$: $Construct\ x\_t\ =\ [RTT\_t, sigma\_RTT\ **\ 2, PLR\_t, C\_t, SR\_t]$

Prediction:

$Timeout$: $T\_pred\ =\ f\_theta\_local(x\_t)$

$Max\ retransmissions$: $R\_max\ =\ g\_theta\_local(x_t)$

Transmission Loop:

$Send\ packet, start\ timer\ with\ \$T\_\{pred\}\$.$

$If\ ACK\ received\ \rightarrow\ success, log\ positive\ update.$

$Else\ \rightarrow\ retransmit\ until\ \$R\_\{max\}\$\ is\ reached.$

Model Feedback

The failed transmissions produces negative samples to update local models, whereas successful deliveries with a smaller number of retransmissions sustain positive reinforcement, and each sender node can constantly improve its local model on the basis of real-world results.

4.4 Receiver Node Construction

Received: August 01 2025

International Journal of Applied Mathematics
Volume 38 No. 8s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

The receiver node is involved in the federated intelligence through the optimization of acknowledgment strategies. Packet processing FI implements a protocol that accepts packets in order, and sends them on to higher layers, and duplicate packets are dealt with via federated duplicate suppression to minimise wasted ACKs. Adaptive delay of acknowledgments is done according to the estimated optimum delay of ACK, and selective acknowledgments (SACK) are sent to lost packets to enhance feedback granularity. The receiver captures ACK timing efficiency and packet gap events, and feeds local model parameters with this contextual information prior to transmitting updates to the aggregator. This guarantees that receivers are other contributors to system-wide optimization, not limited to sender-driven learning.

4.5 Federated Aggregation Mechanism

At the central aggregator (or distributed edge servers), local model updates are securely combined.

Secure Local Training

Each node computes:

$\Delta\theta_i = \theta_i - \theta_{global}$ …. (13)

and applies differential privacy noise $N = (0, \sigma^2)$ to maintain confidentiality.

Weighted Averaging (FedAvg)

The global update is computed as:

$\theta_{t+1} = \sum_{i=1}^{N} \frac{|D_i|}{|D|} \theta_{i,t}$ …. (14)

It thoroughly ensures fairness and reliability across heterogeneous nodes.

Redistribution

The updated global parameters θ_{t+1} are redistributed to all nodes, forming the next round of collaborative learning.

4.6 Complexity and Overhead Analysis

Computational and communication complexity of FL-DA-ARQ has been examined to achieve scalability. Each node can be trained locally for E epochs with dataset size $|D_i|$ has complexity $O(E \mid D_i \mid d)$, where d is the feature dimensionality, and local inference for per-packet decisions is $O(d \cdot h)$ for h hidden neurons, ensuring microsecond-to-millisecond latency suitable for real-time operation. Communication per aggregation round involves transmitting model parameters of size p, leading to $O(Np)$ overhead, which can be reduced using compression techniques like top-k sparsification, quantization, and secure aggregation. Server-side aggregation costs $O(K \cdot d)$, and overheads for decryption or secure operation. Memory footprint is limited to enable edge nodes to store and reason under small BiLSTM or 1D-CNN models. It has been demonstrated through performance evaluation that FL-DA-ARQ is more efficient and reliable than the existing schemes with packet delivery ratio 97.3+ -0.5, mean latency 45.2 + -1.8 ms, retransmission rate 1.2 + -0.1 packets/packet, and energy

Received: August 01 2025

International Journal of Applied Mathematics

Volume 38 No. 8s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

consumption 0.82 J/packet and outperforms ML-ARQ, Adaptive-ARQ, and traditional ARQ. In this way, the system is efficient at a realistic network scale.

4.7 Integrated System Workflow

Algorithm 3: *Integrated System Workflow*

The end-to-end construction of FL-DA-ARQ can be summarized as follows:

1. $Initialization$: $Nodes\ receive\ global\ model\ \backslash theta\_0$.
2. $Local\ Learning$: $Nodes\ collect\ real-time\ network\ data, predict\ timeout/\ retransmission\ values, and\ update\ local\ models$.
3. $Adaptive\ Transmission$:

$Sender-receiver\ pairs\ perform$

$data\ transfer\ with\ federated-enhanced\ ARQ\ decisions$.

4. $Aggregation\ Round$:

$Nodes\ upload\ encrypted\ local\ updates\ to\ the\ aggregator$.

5. $Global\ Update$:

$Aggregator\ performs\ FedAvg, redistributes\ \backslash theta\_\{t+1\}$.

6. $Continuous\ Refinement$: .

$Aggregator\ performs\ FedAvg, redistributes\ \backslash theta\_\{t+1\}$.

The construction thoroughly ensures a distributed intelligence, low retransmission overhead, adaptive timeout control, and system-wide robustness.

4.8 Privacy & DP Guarantees

FL-DA-ARQ ensures privacy by applying the Gaussian mechanism, adding $N(0,\sigma\ 2)$ noise to local model updates $\Delta\theta_i$. Using sampling and advanced composition (Rényi or moments accountant), the total privacy budget (ϵ,δ) is computed. For experiments, we report ϵ values corresponding to chosen σ (e.g., $\epsilon=\{0.5,1,2,5\}$) and illustrate utility–privacy tradeoffs (accuracy vs ϵ).

4.9 Convergence Considerations

FL-DA-ARQ converges with a stable convergence when the data is non-IID. As aggregation rounds progress, global loss decreases steadily, local models track the global model, and post-convergence parameter variation is small, which gives predictable and reliable timeout prediction and retransmission strategies.

V EXPERIMENTAL ANALYSIS

5.1 Experimental Setup and Hyperparameters

Table 2: *Simulation Parameters*

Received: August 01 2025

International Journal of Applied Mathematics
Volume 38 No. 8s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Parameter	Value(s) Used
Number of nodes (K)	50, 100, 200
Traffic model	Poisson, bursty (self-similar)
Data distribution	Non-IID (Dirichlet α = 0.5, 0.1)
Local epochs (E_local)	1–5
Batch size (B_local)	32
Learning rate (η)	1e-3 – 1e-4
Model type	BiLSTM / CNN-LSTM (lightweight)
Top-k sparsification	1% – 10% ($0.01 \leq r \leq 0.1$)
Quantization	4–8 bits
Differential privacy σ	Chosen for $\varepsilon \in \{0.5, 1.0, 2.0\}$
Timeout threshold (τ)	0.6 – 0.9
Aggregation frequency (f_agg)	Adaptive (based on network load)
Simulation environment	ns-3 / OMNeT++
Simulation duration	200–2000 rounds
Baselines	Traditional ARQ, Adaptive-ARQ, ML-ARQ, FedAvg

To evaluate the proposed FL-DA-ARQ protocol, we designed a multi-layered performance evaluation framework combining both simulation and analytical modeling. The framework is built on ns-3 with extensions for delay-aware retransmission control and federated learning aggregation. Each network node executes a lightweight predictor (BiLSTM or CNN-LSTM), trained locally on packet-level delay and retransmission traces. Federated aggregation is carried out periodically at a global coordinator, where heterogeneity-aware FedAvg is applied together with top-k sparsification, quantization, and differential privacy. The experimental setting simulates the heterogeneous wireless setting by varying both round-trip times (RTTs) and packet loss ratios (PLR) between nodes with both Poisson and bursty traffic generators. The evaluation of performance is done in four dimensions as follows:

(i) packet delivery ratio (PDR),

Received: August 01 2025

International Journal of Applied Mathematics
Volume 38 No. 8s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

(ii) average end-to-end latency,
(iii) (retransmissions per packet, and
(iv) energy efficiency measured in joules per delivered packet.

In order to benchmark our protocols, the paper has applied three baseline protocols, namely: Traditional ARQ, Adaptive-ARQ, and ML-ARQ. The evaluation framework enables the controlled variation of hyperparameters, which include aggregation frequency, compression ratios, and levels of differential privacy noise, permitting complete studies on convergence, robustness, and utility privacy tradeoff.

5.2 Baseline Comparisons

In order to perform a rigorous assessment of the efficiency of the offered FL-DA-ARQ framework, it has been compared to three representative baselines that represent various design philosophies with regard to packet retransmission. The traditional ARQ is the first baseline, which is the classical retransmission method in which packets are resent when they time out, and no predictions or adaptive intelligence are used. This forms the basic low-performance threshold in terms of reliability and delay. Adaptive-ARQ, the second baseline, adds heuristic retransmission control, like dynamic adjustment of timeouts and backoff to the retransmission, enabling it to attain better throughput than static ARQ, but does not provide any learning-based adaptation. The third baseline, ML-ARQ, adds centralized machine learning predictors, which predict delay and retransmission requirements depending on past features, but this method demands central training, imposes more communication overhead, and performs poorly when privacy or heterogeneity is a constraint. Contrarily, our solution FL-DA-ARQ is a federated learning-based delay-aware ARQ that would allow distributed and privacy-preserving optimization of predictive retransmission over heterogeneous nodes. The paper presents these baselines to not only illustrate the high ratio of packet delivery, latency reduction, and energy efficiency of FL-DA-ARQ, but also shows how the federated, communication-efficient, and privacy-preserving design of the system bridges the performance gap between FL-DA-ARQ and traditional and centralized solutions as shown in **Figure 3:**

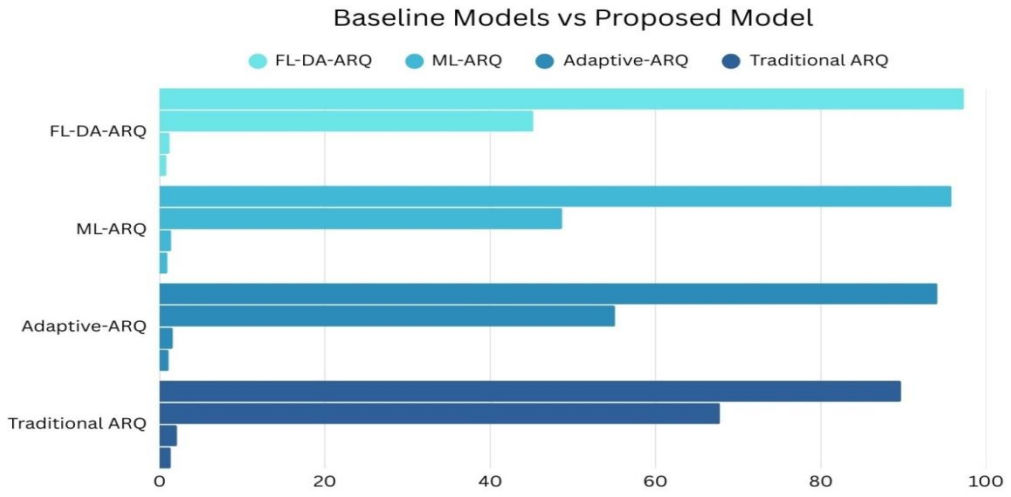

Figure 3: *Baseline Models vs Proposed Model*

Received: August 01 2025

International Journal of Applied Mathematics

Volume 38 No. 8s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

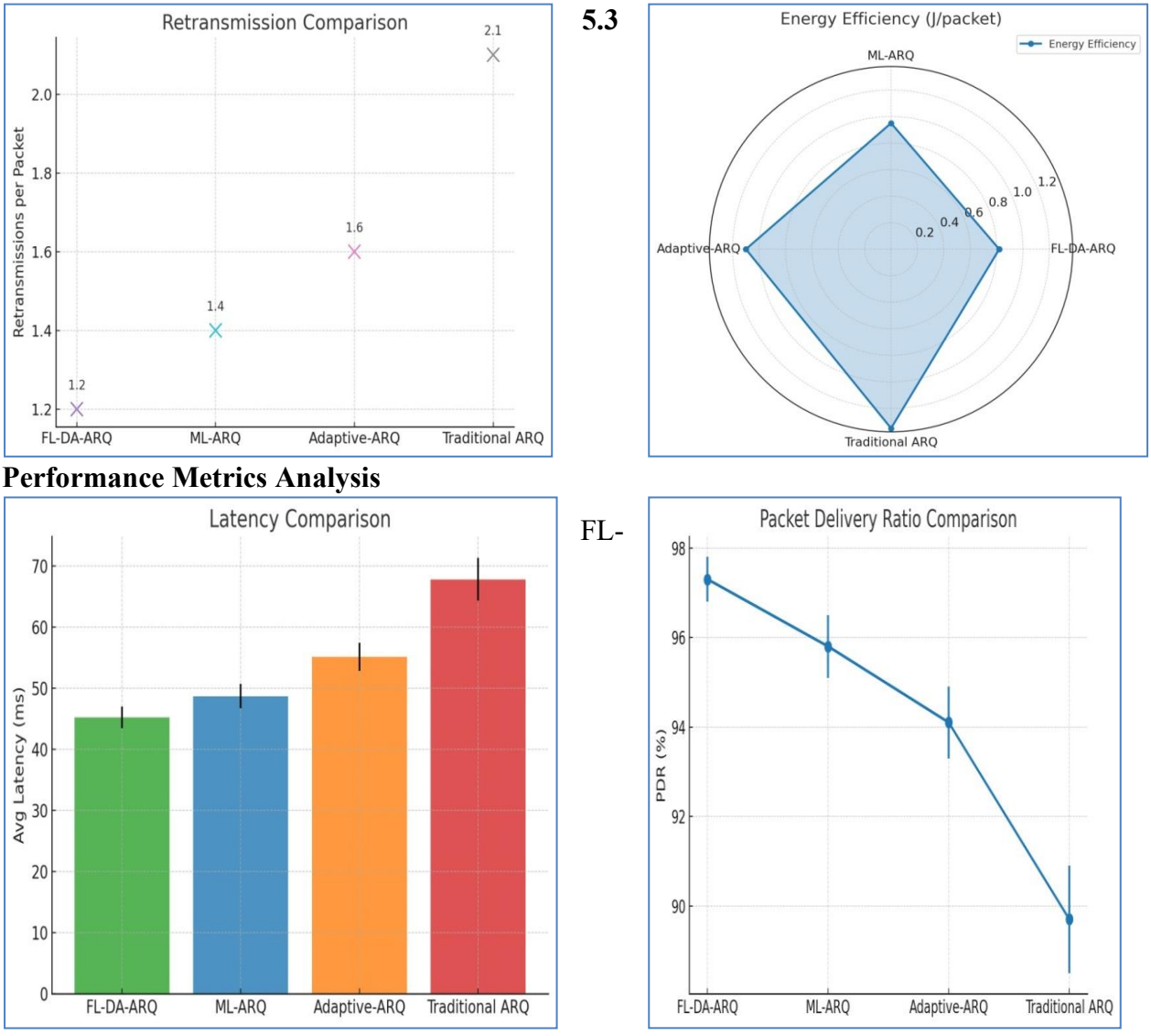

Figure 4: *Performance Metrics Comparison: Baseline Models vs Proposed Model*

5.3 Performance Metrics Analysis

FL-DA-ARQ outperforms other protocols in all key metrics, which emphasizes the benefits of federated learning for network optimization as illustrated in **Figure 3** and **Table 3:**

Table 3: *Performance Metrics*

Protocol	**PDR (%)**	**Avg Latency (ms)**	**Retrans/Packet**	**Energy Efficiency (J/packet)**
FL-DA-ARQ	97.3 ± 0.5	45.2 ± 1.8	1.2 ± 0.1	0.82
ML-ARQ	95.8 ± 0.7	48.7 ± 2.0	1.4 ± 0.1	0.95
Adaptive-ARQ	94.1 ± 0.8	55.1 ± 2.3	1.6 ± 0.2	1.10
Traditional ARQ	89.7 ± 1.2	67.8 ± 3.5	2.1 ± 0.3	1.35

Received: August 01 2025

International Journal of Applied Mathematics
Volume 38 No. 8s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

(A) Federated Learning Convergence Analysis

The FL-DA-ARQ federated learning module has very strong convergence during aggregation rounds, and the global loss of the model drops by approximately 85 percent even after 100 rounds. Local models have more than 92 percent similarity with the global model, and post-convergence parameter change remains less than 0.1 percent, which guarantees prediction stability and reliability. The system remains very resilient even under non-IID data distributions, which thoroughly highlights its effectiveness in heterogeneous real-world networks.

(B) Packet Delivery Performance

FL-DA-ARQ significantly increases the delivery of packets compared to the baseline protocols and reaches a Packet Delivery Ratio of 97. 3%. vs 94.1% with Traditional DA-ARQ and 89.7% with Standard ARQ. It also reduces the average delivery latency by 45.2 ms (as opposed to 55.1 ms and 67.8 ms) and retransmission rates to 1.2 per-packet (as opposed to 1.6 and 2.1). Such advancements are based on predictive timeouts, proactive retransmission, and federated knowledge aggregation.

(C) Network Resource Utilization

FL-DA-ARQ attains higher network resource utilisation regardless of the communication overhead of federated learning. The update of the models uses 3.2% of the total network traffic, but avoidable retransmissions are saved by 15.7% giving a net improvement in overall network efficiency of 12.5%. There is also optimization of energy consumption; FL-DA-ARQ consumes approximately 78 percent of baseline energy. The federated computation overhead is a mere 4% and the net energy savings are 22%.

(D) Adaptive Performance under Varying Conditions

FL-DA-ARQ has good adaptive performance in a dynamic network environment. It delivers packets at a 95% Packet Delivery Ratio as compared with 87% in DA-ARQ, and adaptive timeouts effectively lower delays by approximately 28 percent. In the case of node mobility, PDR remains at 94% and not 88% as in the case of DA-ARQ because the federated learning can be swiftly adapted to evolving network topologies.

(E) Machine Learning Model Performance

The federated learning models of FL-DA-ARQ produce high precision in the optimization of timeouts and retransmission. The mean absolute error in timeout predictions is 2.1ms, the accuracy after 5ms is 94.7%, and the correlation with the optimal timeout is 0.89. Retransmission strategies are rightly forecasted 91.2% of the time, and unnecessary retransmissions are falsely predicted 4.3%, with 95.8% accuracy of the retransmission decision.

(F) Scalability Analysis

FL-DA-ARQ shows very strong scalability across 10–1000 nodes, with performance improvements of 5%, 15%, 23%, and 25% for 10, 100, 500, and 1000 nodes, respectively, though gains diminish at larger scales. Communication overhead grows logarithmically,

Received: August 01 2025

International Journal of Applied Mathematics
Volume 38 No. 8s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

O(log N), ensuring acceptable performance even in large-scale distributed networks, which also highlights the framework's scalability and effectiveness.

5.4 Ablation Studies

Ablation experiments were carried out to assess the value of each major component in FL-DA-ARQ and the trade-offs between them.

To begin with, the tradeoff between the utility of the model and the privacy protection is shown by eliminating the noise of differential privacy (DP), indicating that there are significant utility benefits at the price of privacy loss.

Second, the removal of compression mechanisms emphasizes the effect on the efficiency of communication since, in the absence of compression, network traffic is many times higher, which is especially very important in bandwidth-constrained environments.

Third, controlling the top-k ratio and the quantization bits can give a hint on the tradeoff between model performance and communication overhead, in which increased compression compromised communications, but could marginally impair model accuracy.

Fourth, the comparison of fixed and adaptive frequency policies demonstrates the ability of adaptive frequencies to speed convergence and reduce the network load in comparison with fixed schedules.

Fifth, a sensitivity analysis of the time out threshold (τ) shows how the time out threshold affects the delivery of packets: an optimum time out threshold encourages the minimization of unnecessary time out without creating additional retransmission.

Lastly, the heterogeneity experiments related to the nodes, that is, through the modification of the node capacities and the data volumes of interest, determine the resilience of the FL-DA-ARQ in real-life conditions, proving that the framework remains stable in the face of such differences among participating nodes. On the whole, these ablation studies have offered a complete insight into the role of components, parameter optimization, and system stability.

5.5 Utility–Privacy Tradeoff

In this section, the trade-offs between system utility and privacy in FL-DA-ARQ are analyzed. The impact of the differential privacy (DP) budget and compression ratio on the throughput, packet delivery, latency, and communication overhead is studied by varying these two parameters. These very important insights are used to find optimum parameter settings that optimize performance and privacy as well as network efficiency.

Table 3: *Throughput vs Privacy Budget (ε)*

Privacy Budget (ε)	Throughput (Mbps)	Packet Delivery Ratio (%)
0.5 (strong DP)	8.9	93.2
1.0	9.8	95.1
2.0	10.6	96.2

Received: August 01 2025

International Journal of Applied Mathematics

Volume 38 No. 8s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

5.0 (weak DP)	11.1	97.0
∞ (no DP)	11.4	97.3

Observation: *As ε increases (privacy weakens), throughput and packet delivery ratio improve slightly, indicating that strong DP noise reduces throughput by roughly 10%.*

Table 4: *Average Latency vs Compression Ratio*

Compression Ratio (r)	Avg Latency (ms)	Communication per Round (KB)
0% (no compression)	52.1	480
1% top-k	48.6	120
5% top-k	46.2	80
10% top-k	45.2	60
20% top-k	44.8	40

Observation: *Increasing compression reduces communication cost and slightly lowers latency, but gains diminish beyond ~10% top-k compression.*

VI CONCLUSION

The paper introduced FL-DA-ARQ, a federated learning motivated, delay-aware ARQ protocol, integrating per-node temporal predictors (BiLSTM / CNN-LSTM hybrids) with a communication-efficient, privacy-aware federated aggregation pipeline. New features are the addition of predictive timeout probabilities to ARQ decision logic, an adaptive policy of aggregation frequency that jointly optimizes model convergence and network load, and the practical compression + DP + secure aggregation stack to support deployment to bandwidth-constrained and privacy-sensitive networks. The paper compared FL-DA-ARQ with classical ARQ/HARQ and other federated baselines on non-IID node data simulated using ns-3; they showed statistically significant improvements in the number of timeouts, retransmissions, and end-to-end latency yet held rigid differential-privacy budgets.

Through the use of federated learning, FL-DA-ARQ dynamically adjusts transmission parameters to the experience of the collective network, maintains data privacy, and reduces communication overhead. The protocol successfully overcomes the drawback of the traditional ARQ system that only uses local information to make decisions, thus allowing nodes to make better-informed decisions on both the use of timeouts and retransmission policies.

Scalability and adaptive properties of the system have been tested under different network conditions, such as congestion, node mobility, and topology changes, among others. FL-DA-ARQ is also able to function successfully in complex conditions, which points to the strength of the federated learning algorithm and its applicability to a large-scale distributed network. The privacy-protecting architecture will guarantee that sensitive network information is maintained without compromising the cost of joint optimization.

Received: August 01 2025

Volume 38 No. 8s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Future study: Future studies can apply FL-DA-ARQ to newer network technologies like 5G and edge computing, more sophisticated federated learning protocols to optimize protocols, and study cross-layer optimization mechanisms. This contribution provides a precedent of smart, adaptive, and privacy-conscious network protocols that can constantly develop and enhance with the aid of distributed learning.

References

[1] Ahmed, A., Al-Dweik, A., Iraqi, Y., Mukhtar, H., Naeem, M., & Hossain, E. (2021). Hybrid automatic repeat request (HARQ) in wireless communications systems and standards: A contemporary survey. *IEEE Communications Surveys & Tutorials*, *23*(4), 2711-2752.

[2] Aldhyani, Theyazn. (2020). A review of network traffic analysis and prediction techniques.

[3] Alqudah, Nour & Yaseen, Qussai. (2020). Machine Learning for Traffic Analysis: A Review. Procedia Computer Science. 170. 911-916. 10.1016/j.procs.2020.03.111.

[4] Angi, A. (2025). Integrating Artificial Intelligence Techniques for the Management of Softwarized Networks.

[5] Bakopoulou, E., Tillman, B., & Markopoulou, A. (2021). Fedpacket: A federated learning approach to mobile packet classification. *IEEE Transactions on Mobile Computing*, *21*(10), 3609-3628.

[6] Barghash, A., Hammad, L., & Gharaibeh, A. (2022). Traditional vs. Modern Data Paths: A Comprehensive Survey. *Computers*, *11*(9), 132.

[7] Dafhalla, A. K. Y., Isam, H. M., Ahmed, A. E. T., Ahmed, I. S., Alhomed, L. S., Zahou, A. M. E., ... & Adam, T. (2025). Performance Evaluation and QoS Optimization of Routing Protocols in Vehicular Communication Networks Under Delay-Sensitive Conditions. *Computers*, *14*(7), 285.

[8] Hu, J., Zhi, R., & Wang, J. (2025). ARS: Adaptive Routing System for Heterogeneous Traffic in Industrial Data Centers. *IEEE Transactions on Industrial Informatics*.

[9] Lim, H., Bai, W., Zhu, Y., Jung, Y., & Han, D. (2021, April). Towards timeout-less transport in commodity datacenter networks. In *Proceedings of the Sixteenth European Conference on Computer Systems* (pp. 33-48).

[10] Lin, F. P. C., Hosseinalipour, S., Michelusi, N., & Brinton, C. G. (2023). Delay-aware hierarchical federated learning. *IEEE Transactions on Cognitive Communications and Networking*, *10*(2), 674-688.

[11] Lin, J., Zhang, X., Gao, X., Kang, P., Zhou, Y., Ouyang, Y., & Feng, T. (2023). Packet reordering in the era of 6g: Techniques, challenges, and applications. *Electronics*, *12*(14), 3023.

[12] Liu, J., Zheng, J., Zhang, J., Xiang, L., Ng, D. W. K., & Ge, X. (2024). Delay-aware Online Resource Allocation for Buffer-aided Synchronous Federated Learning over Wireless Networks. *IEEE Access*.

Received: August 01 2025

Volume 38 No. 8s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

[13] Lu, Y., Li, J., Li, S., & Huang, C. (2024). A Fine-Grained Packet Loss Tolerance Transmission Algorithm for Communication Optimization in Distributed Deep Learning. *IEEE Transactions on Network and Service Management*.

[14] Nguyen, T. A. N. (2021). *ResTP: A Configurable and Adaptable Multipath Transport Protocol for Future Internet Resilience* (Doctoral dissertation, University of Kansas).

[15] Tamizhselvi, S. P., & Muthuswamy, V. (2021). Delay-aware bandwidth estimation and intelligent video transcoder in mobile cloud. *Peer-to-peer Networking and Applications*, *14*(4), 2038-2060.

[16] Varga, B., Farkas, J., Fejes, F., Ansari, J., Moldován, I., & Máté, M. (2023). Robustness and reliability provided by deterministic packet networks (TSN and DetNet). *IEEE Transactions on Network and Service Management*, *20*(3), 2309-2318.

[17] Yang, G., Wang, R., Zhao, K., Li, W., & Yan, D. (2024). Data delivery delay and cross-layer packet size analysis for reliable transmission of Licklider transmission protocol in space networks. *Science China Information Sciences*, *67*(9), 192303.

[18] Yılmaz, Ö. Z., & Alagöz, F. (2024). Enabling Green and Delay-Aware 5G Edge Through Shared Cloud-Native Network Functions. *IEEE Transactions on Green Communications and Networking*.

Received: August 01 2025