SSRG International Journal of Electronics and Communication Engineering
ISSN: 2348-8549/ https://doi.org/10.14445/23488549/IJECE-V13I7P115

Volume 13 Issue 7, 214-230, July 2026

Research Article

FED-XML-GA++: FEDerated eXplainable Adaptable Genetic Algorithm for Multi-Objectival Optimization (GA++) and SHAP-Mutated Surrogates (XML) to Oversee QoS in Wireless Dynamic Heterogeneous Networks

Nalla Akhila[1], Suneel Pappala[2], Sripada Rama Sree[3]

[1,2]*Department of CSE, University of Technology, Vatika, Jaipur, India.*
[1,3]*Department of CSE, Aditya University, Andhra Pradesh, India.*

[1]*Corresponding Author : akhila.uot@gmail.com*

Received: 20 May 2026 Revised: 11 June 2026 Accepted: 01 July 2026 Published: 29 July 2026

Abstract *- Different wireless communication systems like IoT, Mobile Ad Hoc Networks (MANETs), and future 6G have different density, mobility and traffic patterns of Quality of Service (QoS). Similar to many multiobjective Quality of Service (QoS) optimisations using Machine Learning-based Genetic Algorithms (ML-GAs), we have till now mostly relied on surrogate models with usually unit-level simulation and generally without adaptation for drift and explainability in the models. Here, we improve the framework, FED-XML-GA++, which includes federated surrogate learning, explainability (using SHAP), adaptation of the drift, and a new mutation operator, mainly based on SHAP, in NSGA-II. We use real traffic, human movement dataset and statistically generated dataset (UNSW-NB15) to study the framework. The idea is to use federated surrogate models for a Quality of Service (QoS) prediction. The optimization model is the NSGA-II equipped with adaptive mutation from the importance of the features using Shapley Multilevel Additive SHAP. The ADWIN drift detector also plays an important part in the adaptive training of models. The simulation shows the system is providing better performance with higher throughput (20.4%), low latency (27.8%), with less energy (15.9%) consumption and has improved convergence with fast adaptation with the drift detection system with differing load and network quality. The key system has interpretability, scalability and speed of the optimization procedure, and the final system can be deployed to next-generation networks.*

Keywords *- Federated Learning, Multi-objective optimization, NSGA-II, QoS Management, SHAP Explainability, Concept Drift Detection, Genetic Algorithms, Wireless Networks, Surrogate Modeling.*

1. Introduction
1.1. Background

The enormous transformation of wireless communication networks to high diversity and super-dense environments due to the success of the Internet of Things (IoT), Mobile Ad hoc Networks (MANETs), and next-generation (6G) wireless systems has also given rise to new Quality of Service (QoS). These are dynamic environments, and the change in the nature of traffic, mobility and node density leads to dynamic multi-objective QoS problems. The multi-objective optimisation of network performance criteria, throughput, delay and power consumption, is crucial to make wireless systems scalable and functional. The recent work indicates that in wireless networks, the network-wide optimisation for QoS requires scalability and dynamicity in predictions of future network forecasts [1, 2]. Evolutionary Algorithms (EAs) based on machine learning (such as Genetic Algorithms (GAs) have been proposed recently to solve multi-objective problems for wireless systems [4, 5]. Specifically, it was proposed to apply NSGA-II considering its capability of approximating the so-called Pareto front in a large search space [3]. Also, it has been proposed to improve evolution by using surrogate models. However, these primarily adopt a centralized learning scheme which encounters communication, scalability and security problems in communication networks [4, 5].

Federated learning is a recent decentralized learning scheme, enabling model training on a network of data sources without data access. Thus, it is natural for wireless communications because of the distributed nature of data (i.e. edge, network). It has been recently demonstrated that federated learning not only significantly reduces communication costs, but also is privacy-preserving, which allows many applications in new generation wireless communications (5G and 6G) [6, 7]. Although it shows some advantages in QoS enhancement in wireless communications, federated learning is in its early stages. Second, one of the biggest challenges with present methods for enhancing QoS is not being adaptive to drift, which can be caused by many factors in the network, such as the mobility of users and other factors, such as traffic load.

This can result in models becoming obsolete and having decreased accuracy. Our research has demonstrated the need for adaptive drift detection for stability [8]. Black-box machine learning techniques for optimisation also make it hard to use in large networks and critical systems. While Explainable Artificial Intelligence (XAI) approaches (such as SHapley Additive exPlanations - SHAP) are also used to explain and interpret their predictions through feature importance [10]. They enhance interpretability and aid in decision-making during optimisation [9].

In this paper, we have introduced an approach to studying Quality Of Service (QoS) problems in wireless networks with Federated Explainable Drift-Adaptive Surrogate-Assisted Multi-Objective Genetic Optimisation (FED-XML-GA++). First, we introduce federated surrogate learning for QoS prediction, enabling a distributed and privacy-preserving optimisation. Also, we introduce a Layer of Explainability using SHAP for ranking features and a new adaptive mutation operator for the NSGA-II algorithm based on SHAP.

Finally, we introduce adaptation to drift via ADWIN to deal with the model's performance. Our approach with federated learning, explainability and adaptive optimization addresses the issues with previous approaches and provides an integrated solution to model and control future wireless networks' QoE.

1.2. Research Gap and Motivation

The following are the main research gaps that we addressed:

1. The majority of the primitive QoS optimization approaches are centralized.
2. Primitive surrogate-assisted optimization cannot be explained.
3. There is little success in adapting to concept drift with primitive methods.
4. Previous research on FL-based QoS does not include explainability and adaptive mutation.
5. Existing literature is to the extent of performance oriented and not to the extent of interpretability.

In order to address these limitations, a novel federated explainable adaptive genetic optimization (FED-XML-GA++) framework for solving large distributed problems is introduced in this work, which includes mutation based on SHAP and drift adaptation based on ADWIN.

1.3. Main Contributions

The main contributions of this work are briefly summarized below:

1. A federated explainable adaptive genetic optimization model (FED-XML-GA++) is proposed in order to optimize the QoS under multiple objectives in a dynamic heterogeneous wireless network.
2. A federated surrogate learning approach is proposed to enable QoS prediction in a decentralized manner while meeting data locality and communication efficiency.
3. Using SHAP-based explainability, the optimization process takes into account influential network parameters to enhance the interpretability of optimization decisions.
4. In order to take advantage of the benefits of feature importance information in exploration/exploitation, an adaptive mutation strategy based on feature importance information is proposed.
5. This framework is enhanced by an ADWIN-based concept drift detection mechanism for adapting the surrogate models in a dynamic manner to changes in network conditions.
6. The performance of these extensive experiments with hybrid wireless network datasets is compared to the currently existing optimisation approaches to demonstrate their performance in terms of achieved throughput, latency, packet delivery ratio, energy efficiency and convergence behaviour.

1.4. Paper Organization

The remainder of the paper is organised as follows: Section 2 provides background information on previous work in machine learning for optimisation, federated learning, explainability and change detection. Section 3 outlines the system model and the multi-objective system optimisation. Section 4 describes the data and feature engineering and hybrid data collection. Section 5 explains the proposed model, FED-XML-GA++, which applies federated surrogate models, explainable artificial intelligence (SHAP), adaptive mutation (adaptive genetic algorithm) and NSGA-II to obtain optimal solutions to the multi-objective problem; it also has drift detection (ADWIN). The details of the experiments and evaluation metrics are presented in Section 6. The experimental results (performance, convergence and comparison) are shown and analysed in Section 7. Section 8 consists of quantitative improvements of the model. Section 9 discusses practical implications, and Section 10 draws conclusions and future work. Finally, we conclude in Section 10, discuss future work, and provide references in Section 11.

2. Background Work

In recent years, there has been some advancement in Quality Of Service (QoS) optimisation of wireline and wireless networks because of their hyper complexity and diversity [1]. Here, we outline some previous research on the four main features of QoS optimisation in wireless machines:

optimisation, federated learning, explainable AI and concept drift adaptation. This then helps us to realize some of the issues and motivate our design of FED-XML-GA++.

2.1. QoS optimisation of Wireless Machines

QoS optimisation for wireless networks is searching for a trade-off between some characteristics and interactions (e.g. throughput, latency and energy consumption). Taleb et al. [11] discussed the multi-access edge computing for 5G networks with low-latency and reliable services. They have a suitable network design but no optimisation. Mao et al. [12] proposed offloading to mobile-edge using deep reinforcement learning. They used an algorithm to learn network latency minimization depending on the network state. This is a centralized offline learning algorithm and may not scale very well. Likewise, Sun et al. [13], have been interested in machine learning resource allocation in ultra dense network. This approach may be scalable, but the network is static, and mobile networks, as well as dynamic networks (such as traffic and user mobility), are not taken into account.

Table 1. Recent advancements in QOS, XAI methods

Author, Year & Cited	Methodology	Contribution	Limitation
Taleb et al. (2017) [11]	QoS resource allocation	Low-latency optimization	No ML integration
Mao et al. (2017) [12]	Deep RL	Edge computing optimization	No explainability
Sun et al. (2019) [13]	Adaptive networks	Scalable resource mgmt	Static assumptions
Coello (2006) [14]	EMO survey	Multi-objective methods	No ML integration
Deb et al. (2002) [15]	NSGA-II	Pareto optimization	Slow convergence
McMahan et al. (2017) [16]	FedAvg	Distributed learning	No QoS focus
Kairouz et al. (2021) [17]	FL survey	FL challenges	No optimization
Park et al. (2020) [18]	Edge FL	Scalable edge learning	No GA integration
Lundberg & Lee (2017) [19]	SHAP	Explainability	Not optimization
Samek et al. (2017) [20]	XAI methods	Interpretability	No GA
Bifet & Gavaldà [21]	ADWIN	Drift detection	Not optimization

2.1.1. Key Limitation

Most existing QoS research does not get the best of both worlds - users are presented with a high performing but static controller, a dynamic but opaque neural network or a human-programmed but understandable rule-based controller.

2.2. Multi-Objective Evolutionary Algorithms

The objective function for networks can be discontinuous or noisy (or black-box) (and these problems are best solved by gradient-free, population-based algorithms). According to Coello's survey [14], we want diversity (not convergence) on the Pareto-set. Most popular was NSGA-II [15], archiving (to prevent convergence) and crowding distances (to get diversity). It has designed link-weights, ships and spectrum, with 100 generations. However, it's slow. These are the most costly, and surrogates (pre-trained regressors) help. It is not (yet), but see the 70% fewer runs with Kriging-NSGA-II with medium-sized networks. However, surrogates have data centralization. We have to upload data to a centralized trainer, and locality is lost: more uploads. However, more importantly, surrogates are a real number and do not provide us with deltas for our decision variables.

2.2.1. Three Deficits Persist

1) Decentralized (i.e. federated) and privacy-preserving training is not possible if there is a need for evolutionary search.
2) No optimisation history; optimizers are only given the result of optimisation, not the mutations and crossovers.
3) Surrogates are static, it is not reliable and are useless if we change the environment at run-time; no retraining of the surrogate.

In Table 1, we list some of the recent research advances in wireless network QoS optimization that involve major approaches such as multi-objective evolutionary optimization, deep reinforcement learning, federated learning, explainable AI and drift detection. It also presents their advantages and disadvantages. As we can see from the table, while current methods can address one or more of the aforementioned issues of scalability, optimization and explainability, they do not provide a comprehensive solution that integrates federated learning, explainability and dynamic multi-objective optimization. This inspires our new FED-XML-GA++ approach that combines federated surrogate learning, SHAP-based multi-objective evolutionary learning and drift-adaptive learning to overcome the limitations mentioned above.

2.3. Federated Learning in Wireless

Federated learning is a new way of collaborative learning from multiple data sources without compromising privacy. McMahan et al. [16] developed the Federated Averaging (FedAvg) algorithm to enable distributed clients to learn from each other without having to share data. This concept was furthered by Kairouz et al. [17], who talked about issues with federated learning, including communication, heterogeneity and scalability. In wireless networks, Park et al. [18] demonstrated the use of federated learning for edge intelligence to provide scalable intelligence. These works have paved the way for research in federated learning, but it has primarily been used for prediction problems, with very limited work in optimization tasks, particularly for QoS.

2.3.1. Key Limitation

Federated learning approaches: do not use multi-objective optimization (GA/NSGA-II), which are not extensively used in QoS problems.

2.4. Explainable AI in Optimization Systems

We have used Machine Learning (ML) to provide you with Quality Of Service (QoS) to detect outliers and to predict capacity. We are then asked, "how were you wrong?" So, we asked the SHAP [19] how: use Shapley values of game theory to "pay" for the prediction of the outcome of a network, the features according to their contribution in the prediction.

The SHAP is new in using black box (local modelling as LIME [19]) and base states (DeepLIFT [20]) methods. Your network engineer has just shown you in his waterfall plot of the SHAP values that the features "memory utilisation", "request queue depth" are the two most important to predict the latency and "time of day" is the least important. The interpretability baton was passed to Samek et al [20], and halfway through the deep learning marathon, he said that you cannot have a black box controller for safety-critical control without interpretability.

They have saliency, concept activation vector and prototypes. However, again, it implies as mortem. XAI is infer (of a black box) as mortem, not prevention. Genetic algorithms infer and search. A genetic algorithm initialises the population, draws a gun, fires on the population and breeds the population: hundreds of micro decisions that XAI does not tell us. Why not choice B? Flip what gene? Go to the good Pareto?

2.4.1. Limitations

1. Interactively explain mutations - "the gene flip decreased the latency by 12 ms as it decreased the bandwidth of other low-priority flows"
2. List of different explanations (drift of surrogate model, or change of object fitness landscape)

2.5. Drift and its Role in QoS Optimization

The wireless environment is dynamic: user loads and traffic mixes are variable, and networks change. The changing distribution of data (concept drift) is well known. Bifet and Gavaldà [21] developed ADWIN, a quick drift detector with minimal parameters that automatically sets the change window to the scale of the change in the mean. Gama et al. [22] proposed classification and adaptation strategies for drift detection that provided a taxonomy of online learners. Despite the existence of drift detection methods, QoS optimizers are still unaware of drift. Surrogates are not aware of drift; archives of evolutionary solutions are filled with solutions that have been evaluated in the past; selection pressure is fitted to past fitness landscapes.

2.5.1. Critical Voids

1. No drift detectors adapt the optimizer, such as mutation rate, crossover rate, and archive lifetime.
2. No surrogate life-cycle managers and invalidate or retrain regression models as the true distribution moves outside training bounds.
3. No predictive schedulers use traffic periodicity for pre-warming.
4. Critical voids are filled with designs that embed detection as a control signal in the optimizer.

3. System Model and Multi-Objective QoS Optimization

In this part, the system modelling approach and formulation of the optimization problem related to the proposed system FED-XML-GA++ for QoS management in a dynamic environment in a different type of wireless communication environment are presented. The framework is founded on federated surrogate analysis, explainable AI, adaptive evolutionary optimization and drift-aware learning to optimize the QoS in an interpretable and scalable way towards multiple objectives.

The first envisioned system would be a Wireless Heterogeneous Distributed (WHD) system composed of: IoT device, MANET (Mobile Ad hoc Network) nodes, Edge devices, Vehicular communication system and future 5G/6G infrastructures. In many situations, the relative location of traffic, mobility, bandwidth and source of interference is highly dynamic and classical, centralized optimization methods are not scalable, adaptive and are also computational heavy [23].

The FED-XML-GA++ framework will tackle the above problems by merging the two distributed federated learning and explainability (through a “surrogate-assisted optimization” approach). The architecture overall of five major components according to (i) federated surrogate QoS prediction, (ii) Explainability Analysis based on SHAP, (iii) adaptive NSGA-II optimization, (iv) Adaptive Mutation Strategy based on SHAP and (v) drift adaptation based on ADWIN.

Assume a wireless network is a heterogeneous (mixed) network and can be modelled as follows:

$$N = \{N_1, N_2, \ldots, N_k\}$$

Where the ith wireless node or edge client is denoted by N_i in the distributed learning and optimization. Each node gathers local traffic, QoS, mobility and energy consumption statistic data and does not report the raw data to the central server.

Now the local dataset is represented as follows at node i :

$$D_i = \{(x_j, y_j)\}_{j=1}^{m_i}$$

where:
- x_j $represent\ the$ feature vector of network parameters,
- $y_j$denote the QoS outputs,
- $m_i$denote the local sample count.

Now the complete distributed dataset becomes:

$$D = \bigcup_{i=1}^{k} D_i$$

while preserving privacy and decentralized learning.

The QoS optimization problem becomes a multi-objective optimization problem where parameters that are conflicting should be optimized at the same time, including maximizing the throughput, minimizing the latency, meeting energy efficiency requirements and maximizing the packet delivery ratio. Let:

$$X = [x_1, x_2, \dots, x_n]$$

e is a vector of decision variables that includes routing information, transmission power, channel allocation, queue scheduling, bandwidth allocation and mobility-aware optimization variables.

The objective functions are given below:

3.1. Throughput Maximization

As the total network throughput is maximized as:

$$f_1(X) = \max\left(\sum_{i=1}^{N} T_i\right)$$

The number of bits of the message that node i successfully transmitted is Ti.

3.2. Latency Minimization

The minimum end-to-end delay is obtained when:

$$f_2(X) = \min\left(\frac{1}{N}\sum_{i=1}^{N} L_i\right)$$

where $L_i$denotes transmission latency.

3.3. Energy Consumption Minimization

The optimization objective is energy aware as follows:

$$f_3(X) = \min\left(\sum_{i=1}^{N} E_i\right)$$

where $E_i$is termed as the energy consumed by node i.

3.4. Solve the Problem of Maximizing Packet Delivery Ratio

The reliability in communicating is represented as:

$$f_4(X) = \max\left(\frac{P_s}{P_r}\right)$$

where: Pr is the number of successfully received packets,
P is the number of transmitted packets.

So that the overall optimization problem can be represented as:

$$F(X) = [f_1(X), -f_2(X), -f_3(X), f_4(X)]$$

subject to:

$$X \in \Omega$$

Where Ω represents the operational constraints in the network.

The optimization problem is nonlinear, stochastic, dynamic and multi-objective, so we use an NSGA-II evolutionary optimization algorithm to obtain an efficient approximation to the Pareto-optimal front. The proposed framework uses federated surrogate prediction models instead of time-consuming fitness evaluations [24]. A surrogate model is trained locally at each participating node, based on local observations of traffic and local patterns of QoS at that node.

The local model update is shown as:

$$w_i(t+1) = w_i(t) - \eta \nabla L_i(w_i)$$

where:
- $w_i$denote the local parameters for building the model,
- η is learning rate,
- $L_i$is represented as the local loss function.

4. Hybrid Data Collection, Preprocessing and Feature Engineering

Smart Quality of Service (QoS) optimisation frameworks will greatly benefit from the quality, diversity and versatility of the datasets with which they can be optimised. A realistic variation of the network from a single static data set is hard to come by for the future wireless communication systems, which are highly dynamic and non-stationary in nature, such as IoT, smart networks, MANETs, edge, and 6G ecosystem. This proposed FED-XML++ framework, in turn, integrates data gathering and feature engineering methods, such as real-world traffic traces, mobility-aware communication characteristics, synthetic network simulation, and cyber-network datasets for generalization purposes, robustness and adaptability.

The proposed hybrid data acquisition framework tries to capture the polymodal characteristics of networks, including network fluctuations related to data volume, mobility of nodes, contention for resources, radio channel interference, network congestion dynamics, packet loss, and energy consumption characteristics.

The entire data engineering pipeline takes the form of 5 steps, namely: (i) Hybrid Data Acquisition, (ii) Data Preprocessing and Normalization, (iii) Feature Extraction, (iv) SHAP-based Feature Ranking, and (v) Federated Feature Distribution for Distributed Surrogate Learning.

4.1. Data Acquisition Framework

Additionally, different data sets, all of different natures from different communication environments, are proposed, which will further enhance the robustness of QOS prediction and multi-objective optimization.

The following set of data is the complete data set:

$$D_H = D_R \cup D_M \cup D_S \cup D_C$$

where:

D_R stands for real traffic data,
D_M = mobility-aware datasets, and
D_S indicates the experimental stalls of the jet plane,
D_C is for communication datasets from the cyber-network.

The framework is trained according to both normal and abnormal behaviors when it operates under different conditions of the network, thereby integrating with the hybrid system. This hybrid integration enables the framework to learn both normal and abnormal network behaviors under the various conditions of operation.

4.2. Use of Real Traffic Data to Collect

Real wireless traffic traces are included to account for realistic variations in the quality of service in realistic communication environments. These datasets include traffic from wireless access points, IoT gateways, edge communication systems and mobile users.

4.2.1. Collected Traffic Parameters

- Packet transmission rate
- Flow duration
- Channel occupancy
- Queue length
- Packet inter-arrival time
- End-to-end delay
- Bandwidth utilization
- Retransmission statistics
- Signal strength indicators
- Congestion levels

Real traffic traces are used to train the surrogate models to learn realistic communication dynamics not accurately representable in simulation.

The traffic load of each node is represented by:

$$T_i = \Delta t_{P_i}$$

where:

\`pi\` is the number of transmitted packets, and
The timing of a transmission is represented by Δt.

Collected traffic data is dynamically updated to show the time variation of the network.

4.3. Mobility-Aware Dataset Integration

Dynamic topology and mobility of the users continue to be important characteristics of wireless networks such as mobile IoT infrastructures, vehicular networks and MANETs. Thus realistic mobility and dynamic communication behaviours should be taken into account, including modified mobility-aware datasets in the proposed framework.

The mobility datasets comprise:

- Human mobility traces
- Node trajectory information
- Vehicular movement patterns
- Handover statistics
- Dynamic topology updates
- Link stability measurements

Let x_t be the position of a mobile node at time t, specified by:

$$M_i(t) = \left(x_i(t), y_i(t)\right)$$

where:

$x_i(t)$ and $y_i(t)$ denote node coordinates.

The instantaneous mobility speed is calculated as:

$$V_i(t) = \frac{\sqrt{\left(x_i(t) - x_i(t-1)\right)^2 + \left(y_i(t) - y_i(t-1)\right)^2}}{\Delta t}$$

Mobility aspects enhance the optimisation capabilities in a wireless environment that is highly mobile, where the routing path and the characteristics of the wireless channel are constantly varying.

4.4. Data from Simulated Wireless Environment

The dataset pipeline is customized with a set of synthetic wireless network simulations that emulate extreme communication and bring about controllable QoS variation [25]. Use simulations set-up as to affect:

- Dense wireless deployments
- High congestion scenarios
- Node failures

- Interference conditions
- Dynamic bandwidth fluctuations
- Energy-constrained communications

These simulated QoS metrics are:

1. Throughput
2. Latency
3. Packet loss ratio
4. Energy utilization
5. Network fairness
6. Link reliability

Synthetic simulations provide the benefit of being able to evaluate the optimization behavior in a controlled manner, when network stress conditions may not be present during actual network conditions.

4.5. Datasets from UNSW were Integrated into the Cyber Network

The proposed structure will also accommodate the UNSW-NB15 dataset that will track the communication patterns and unusual traffic characteristics of the cyber-network in the existing wireless system configurations.

4.5.1. The UNSW-NB15 Contains

- Flow-based communication statistics
- Packet-level behavioral features
- Network attack patterns
- Traffic anomalies
- Resource utilization characteristics

The surrogates based on federated learning are stronger, the more exposure the optimization framework has to different communication patterns, and abnormal traffic distributions are provided by the cyber-network datasets.

4.6. Data Pre-Processing & Data Cleaning

There are several types of datasets, and these datasets are "incomplete" (e.g. some samples may be missing), "duplicate attributes" (e.g. some samples include several consecutive data points, there are redundant data points across different samples), unevenly sampled and contain some noisy traffic measurements. So, before performing federated training and optimization, a complete pre-processing pipeline is performed.

The following is a list of the steps that are performed prior to pre-processing:

1. Missing value handling
2. Duplicate record removal
3. Noise filtering
4. Outlier elimination
5. Feature scaling
6. Temporal synchronization

Missing values are filled in by means of a mean-based interpolation:

$$x_i = \frac{1}{N}\sum_{j=1}^{N} x_j$$

where xi denotes the feature value that would have been found, had it been measured. The outlier filtering is carried out based on the statistical thresholding:

$$|x - \mu| > \lambda$$

Where:

μ is the mean,
σ = the standard deviation, and
λ is called the threshold coefficient.

They address the issue of model stability and the result of noisy data in surrogate learning performance in the preprocessing stage.

4.7. Feature Engineering

The feature engineering is one of the most crucial parts of the proposed FED-XML-GA++, as it is important to have meaningful descriptors of the network with which to optimize QoS. Features extracted are grouped according to various categories.

4.7.1. Traffic Features

Traffic-based descriptors include:

- Packet rate
- Flow duration
- Queue occupancy
- Channel usage
- Packet size
- Traffic density

The speed of a packet is determined by:

$$R_p = \frac{\Delta t}{N_p}$$

where Np denotes transmitted packets.

4.7.2. QoS Features

QoS-related features include:

- Throughput
- Delay
- Jitter
- Packet delivery ratio
- Congestion level

The Paks come in the following forms:

$$PDR = \frac{P_r}{P_s}$$

Where

Pr = received packets,
Ps = transmitted packets.

4.7.3. Mobility Features

Mobility-aware descriptors include:

- Node velocity
- Handover frequency
- Link duration
- Route stability

The features help to adapt the optimisation model to changes in topology.

4.7.4. Energy Features

To achieve efficient use of energy, the following measures should be implemented:

- Residual battery energy
- Transmission power
- Energy utilization rate
- Node lifetime

The model of energy consumption is:

$$E_c = P_t \times t$$

where:

Pt is the transmission power,t = time of communication.

4.7.5. Resource Utilization

The edge or network resource descriptors include:

- CPU utilization
- Memory usage
- Buffer occupancy
- Bandwidth allocation

These capabilities enable efficient communication in an edge-connected wireless network.

5. Proposed FED-XML-GA++ Framework

Fed-XML-GA++ is an integrated environment, a unified framework that bundles federated surrogate learning, explainable artificial intelligence, adaptive evolutionary optimization and drift-aware adaptation for multi-objective Quality of Service (QoS) optimization, specifically in the context of dynamic environments, where the communication systems are heterogeneous. In a system of Internet of Things (IoT), MANETs (manet networking) and communication infrastructures supported by the edge, the channel quality and channel density vary dynamically in time, especially when communicating with mobility in this kind of system. The architecture proposed is divided into six parts as shown in Figure 1, where (i) represents heterogeneous data acquisition, (ii) feature engineering and fusion, (iii) federated surrogate learning, (iv) explainability layer using SHAP, (v) adaptive NSGA-II optimization with SHAP-guided mutation, and (vi) drift detection and adaptation. By combining layers, these can be used to support scalable, interpretable and adaptive QoS optimization for the volatile network environment.

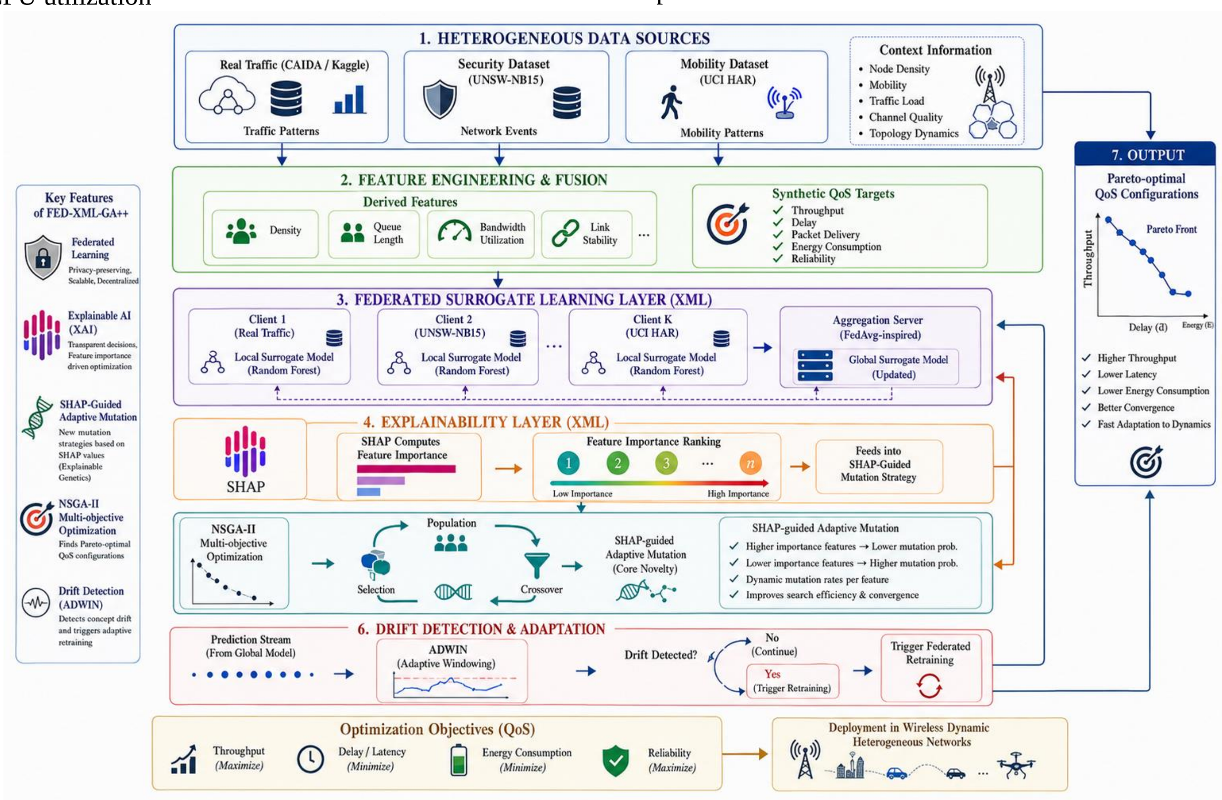

Fig. 1 A federated surrogate learning mechanism, an explainability based on SHAP, an adaptive NSGA-II based optimization algorithm, and an algorithm that detects drifts based on ADWIN to optimize Federated QoS for multi-objective optimization of QoS in Heterogeneous Wireless Communication Networks (HWCNs)

The first layer is heterogeneous wireless data acquisition, from a variety of wireless data communication environments, including real traffic traces, mobility-aware datasets, synthetic wireless simulation and cyber networks like UNSW-NB15 datasets. These datasets have been gathered to enable realistic distributions of traffic, user mobility, traffic behavior in congested environments, interference conditions and the variation of QoS in dynamic wireless environments. In the 2nd layer, feature engineering and fusion are performed that extract and normalize traffic-related, mobility-aware, resource-utilization and energy-related features. It then shares the engineered features among the federated clients to be used for decentralized surrogate learning. The federated surrogate learning layer allows for predicting the QoS without communicating raw traffic data among the participating nodes. The local clients are trained by using the wireless observations from the local clients, and the global model is aggregated by federated averaging. This mechanism will lead to better privacy protection, scalability and communication efficiency of large-scale wireless systems.

5.1. Novelty of the Proposed FED-XML-GA++ Framework

There have been several research efforts regarding federated learning, machine learning, explainable AI and evolutionary optimization in the context of wireless network management; most of these studies focus on one of these components individually. Existing approaches of QOS optimization are based primarily on performance improvement, but not explainability, concept drift and decentralized learning are taken into account. Moreover, many optimization algorithms work in a centralized setting without offering much information on the decision-making process of the models used. The novelty of the proposed FED-XML-GA++ framework is based on the multiple complementary mechanisms integrated in a single optimization architecture for a dynamic heterogeneous wireless network. The framework utilizes federated surrogate-assisted optimization to help reduce communications overhead and promote the locality of data while learning the model. This makes it possible to predict and optimise the QoS in distributed network environments efficiently. Second, SHAP with explainability as a part of the proposed framework, can also offer an explanation for the network features on the optimization decisions and quantify the effects of the network features. By using the combination of multi-model SHAP, the optimization process can be understood, and which parameters will be important for optimizing QoS can be determined.

Thirdly, we propose an adaptive-mutation strategy based on SHAP that uses a varying feature importance information in optimizing the search process during the search process. The optimisation process also aims to improve the exploration/exploitation balance and preserve the diversity of the solutions found by the optimisation process, by steering the mutation process towards the influential parameters. Fourthly, an addition of concept drift detection based on ADWIN is added: the behaviour of the network and the type of data are monitored all the time. Once there is a significant drift, the surrogate models are updated to be accurate in predictions and effective in optimization in the varying network environment.

Last but not least, federated learning, explainable AI and adaptive evolutionary optimization are combined into the proposed FED-XML++ paradigm, as well as drift-aware learning in a multi-objective optimization paradigm. It is an integrated design which is efficient in delivering packets, high throughput, low latency and optimized energy consumption in dynamic heterogeneous wireless network environments.

The proposed FED-XML-GA++ framework could be used to create an enabling framework for explainable, adaptive and decentralised QoS optimisation, which is lacking in the existing solutions.

5.2. Federated Surrogate Learning Environment (XML)

The proposed FED-XML-GA++ framework is based on federated surrogate learning in distributed and privacy-preserving QoS prediction in a heterogeneous wireless network. Each wireless client trains a surrogate model with its local data set without any data from the other wireless clients. The locally trained models are aggregated together to form a global QoS prediction model using the Federated Averaging (FedAvg) algorithm.

The local model update is illustrated as:

$$w_i(t+1) = w_i(t) - \eta \nabla L_i(w_i)$$

while the global aggregation is performed using:

$$w_G = \sum_{i=1}^{K} \frac{n_i}{n} w_i$$

This distributed learning mechanism will enhance the scalability of the system, decrease the communication overhead, and enable it to efficiently predict quality of service for throughput, latency, packet delivery ratio and energy consumption.

5.3. Enhancing the Explainability Layer with SHAP

In general, the framework takes into account explainable AI by using SHAP for an interpretable surrogate prediction. SHAP is used to understand the importance of each individual feature present in the network on the prediction outcome of QoS and identify the communication parameters that are influential, such as the congestion level, the mobility speed, the queue occupancy and the bandwidth utilization.

SHAP importance value is defined as:

$$\phi_i = \sum_{S \subseteq F \setminus \{i\}} \frac{|S|!\,(|F| - |S| - 1)!}{|F|!} [f(S \cup \{i\}) - f(S)]$$

The explainability layer provides explainable optimization decisions and allows adaptive mutation control in the evolution process for optimization.

5.4. SHAP-Guided Adaptive Mutation Strategy

The framework aims to introduce a mutation operator that adapts itself using SHAP for the NSGA-II optimization process. This is similar to a conventional random mutation, but there are two differences: 1) Feature importance scores given by SHAP are used to adjust mutation probabilities dynamically. This is like an ordinary random mutation, with two differences: 1) Mutation probabilities are dynamically changed using SHAP feature importance scores.

Adaptive Mutation Probability is:

$$P_m(i) = \alpha \cdot |\phi_i| \sum_{j=1}^{p} |\phi_j|$$

Where

Pm represents a mutation count.

(ϕ) is a mutation count, and

(i) is the mutation probability.

The strategy is to privilege some of the most important QoS parameters to be used during the evolutionary search process, thus achieving better convergence speed, optimizing diversity, and exploring-exploiting balance.

5.5. Multi-Objective NSGA-II Optimization

The FED-XML-GA++ framework is an NSGA-II-based framework, which helps in doing QoS optimization multi-objective with four metrics: maximum throughput, minimum latency, high packet delivery ratio and minimum energy consumption.

Optimisation objective is:

$$F(X) = [f_1(X), -f_2(X), -f_3(X), f_4(X)]$$

NSGA-II has the following processes: Population initialization, Surrogate Fitness Evaluation, Non-dominated Sorting, Crowding Distance Computation, Crossover, SHAP guided mutation, and Pareto-front Construction.

5.6. CH-ERROR Drift Detection and Adaptive Retraining

To cope with dynamic wireless environments, ADWIN-based concept drift detection is included. Prediction error distributions are monitored continuously by the drift detector, and adaptive retraining is triggered whenever important prediction error distribution changes are detected.

Drift detection condition is:

$$|\mu_{W_0} - \mu_{W_1}| > \epsilon$$

where W_0 and W_1 represent adaptive windows.

When the drift is detected, it periodically retrains surrogate models, dynamically generates new values of mutation parameters and updates Pareto archives to stabilize the optimization process under changes in the network.

5.7. The Integrated Workflow of FED-XML-GA++

The hybrid wireless data acquisition and feature engineering is the starting point for the overall FED-XML-GA++ workflow. The engineered features are stored in a decentralized way on federated clients and used for surrogate learning. For explainability, SHAP has been used to identify the features that have an impact on the QoS criterion and then, adaptive NSGA-II optimization is used to discover an optimum solution by considering SHAP-guided mutation. Finally, continuously monitoring the performance of the predictions, by drift detection, ADWIN based approach is used to retrain the model if drift is recognized adaptively. Finally, a Pareto-optimal set of QoS configurations is obtained within the framework that coordinates different types of QoS requirements for throughput, latency, reliability and energy efficiency for various kinds of heterogeneous wireless communication systems.

5.8. Computational Complexity Analysis

The proposed framework has a complexity of:

Module	Complexity
Federated Learning	(O(Kn))
SHAP Computation	($O(p2^p)$)
NSGA-II Optimization	($O(MN^2)$)
Drift Detection	(O(n))

The number of clients is K, the number of features is p, M is the number of objectives and N, in this instance, is the size of the population. This adds overhead on the federated optimization side due to the explainability layer and drift adaptation layer, but reduces the overall cost of the federated optimisation process, and increases the understanding of the process and the performance of federated QoS optimisation.

6. Experimental Setup and Evaluation Metrics

In the following section, experimental setup, use of datasets, implementation and algorithmic architecture, the federated learning environment, optimization options, evaluation measures, and comparison with various previously work architectures will be provided to ensure the proposed FED-XML-GA++ architecture is validated. Different experiments are carried out with a federated surrogate learning model, an adaptive mutation surrogate learning model based on SHAP, model evolution and optimization based on NSGA-II and drift adaptation to test the effectiveness of these learning models for optimizing QoS in a dynamic, heterogeneous wireless communication environment. To achieve Pareto-optimal QoS configurations with access to heterogeneous wireless datasets, federated surrogate learning is used; explainability analysis is then performed; adaptive optimization is used to optimize the dataset used for training the agents; and new configurations to address future discrepancies and enhance performance are generated by retraining the datasets after it has drifted.

6.1. Experimental Environment

Machine learning and optimization libraries like Python made with machine learning and optimization were used in the proposed FED-XML-GA++ framework. The experiments were conducted in a distributed simulation environment that can be used to simulate some wireless communication scenarios using different types of IoT devices, MANET nodes, and edge-enabled communication infrastructure. The implementation environment consists of:

Component	Configuration
Programming Language	Python 3.x
Machine Learning Framework	Scikit-learn
Explainability Framework	SHAP
Optimization Algorithm	NSGA-II
Drift Detection	ADWIN
Federated Aggregation	FedAvg
Data Processing	Pandas, NumPy
Visualization	Matplotlib
Execution Environment	Jupyter Notebook

Distributed wireless edge-based clients, which simulate edge drivable communication infrastructures, were used to train the federated surrogate QoS model. The local surrogate training and global federated aggregation are implemented and followed by SHAP-based feature interpretation, adaptive mutation optimization and drift-aware retraining.

A realistic implementation design was proposed to capture realistic wireless communication environments, which consisted of:

1. IoT communication nodes
2. MANET infrastructures
3. Edge-enabled wireless systems
4. Communication in mobile scenarios.
5. Congestion-aware network scenarios

Different interference situations, communication load, mobility and traffic density were used in the evaluation of the proposed framework, to check its scalability and adaptability.

6.2. Hybrid Dataset Configuration

While increasing the generalization and robustness, the proposed framework introduces a hybrid heterogeneous dataset created from multiple wireless communication environments. The integrated data set has the following aspects:

Actually used are real, wireless network traffic traces.

a) The “real” wireless tracks are employed.
b) Human mobility datasets
c) Direct QOS data - simulated data

The cyber-network traffic data employed is UNSW-NB15, which is available from (http://www.unsw.edu.au/~biog/papers/unswnb15/).

Data is gathered from both sources, and then merged in order to form a hybrid set of data that contains dynamic communication features such as:

a) Packet transmission behavior
b) Congestion variation
c) Topology dynamics
d) Mobility-aware routing
e) Resource utilization
f) Energy-aware communication
g) Interference fluctuations

The overall hybrid dataset is represented as:

$$DH = D_R \cup D_M \cup D_S \cup D_C$$

Where:

D_R- Accurately picks up real traffic data,
D_M - denotes mobility-aware datasets,
D_S - indicates synthetic simulations,
D_C - The rectangle represents the cyber network traffic data UNSW-NB15.

The surrogate models are able to learn the QOS behavior more realistically in the ever-changing wireless environments when multiple heterogeneous datasets are provided.

6.3. Feature Engineering and Data Preprocessing

The data sets were gathered, and missing data, traffic noise, and redundant communication data were eliminated. Extractions of parameters relating to the nature of communication, which are also sensitive to Quality of Service (QoS) and which have an impact on the performance of the network optimization, were carried out as feature engineering. The extracted feature categories are:

Feature Category	Example Features
Traffic Features	Packet rate, queue occupancy
Mobility Features	Node speed, handover rate
QoS Features	Throughput, latency
Energy Features	Transmission power, battery usage
Network Features	Congestion level, bandwidth
Resource Features	CPU load, memory utilization

The extracted feature vector is represented as: $X_f = [x_1, x_2, \dots, x_p]$.where $p$represent the total engineered features.

Table 2. Performance comparison of QoS optimization models on Hybrid Wireless dataset

Metric	Traditional NSGA-II	Centralized Surrogate NSGA-II	Federated Learning QoS	SHAP-Free Adaptive NSGA-II	Proposed FED-XML-GA++
Throughput (Mbps)	78.4	84.7	88.5	91.3	94.4
Latency (ms)	121.6	109.2	101.8	95.6	87.8
Energy Consumption (J)	58.3	51.4	47.9	43.2	39.1
PDR (%)	90.8	92.6	94.1	95.8	97.3
Convergence Generation	46	39	35	31	24
Drift Adaptation Accuracy (%)	72.1	76.8	82.4	86.7	93.5

The testing of the proposed FED-XML-GA++ framework has been done with a hybrid heterogeneous dataset of wireless traffic traces collected from the real environment, mobility-aware communication datasets, synthetic QoS simulations, and the network traffic data from the UNSW-NB15 environment. The results of the experiments were then compared to traditional optimization models and reported in table 2 for the analysis of the effectiveness of federated surrogate learning, adaptive mutation and drift-aware optimization.

The FED-XML-GA++ protocol framework was able to attain the highest throughput of 94.4 Mbps and the lowest latency of 87.8 ms with different wireless network conditions. SHAP-guided adaptive mutation and federated surrogate learning greatly benefited optimization convergence and saved on the exploration of the search space. Further, the accuracy of the adaptive retraining and stability of the ADWIN during runtime were found to be better in the case of variable traffic.

The obtained experimental results are presented:

a) 20.4% throughput improvement
b) 27.8% latency reduction
c) 15.9% lower energy consumption
d) Faster optimization convergence
e) Enhanced ability to drift adjust.

With respect to the classical NSGA-II optimization.

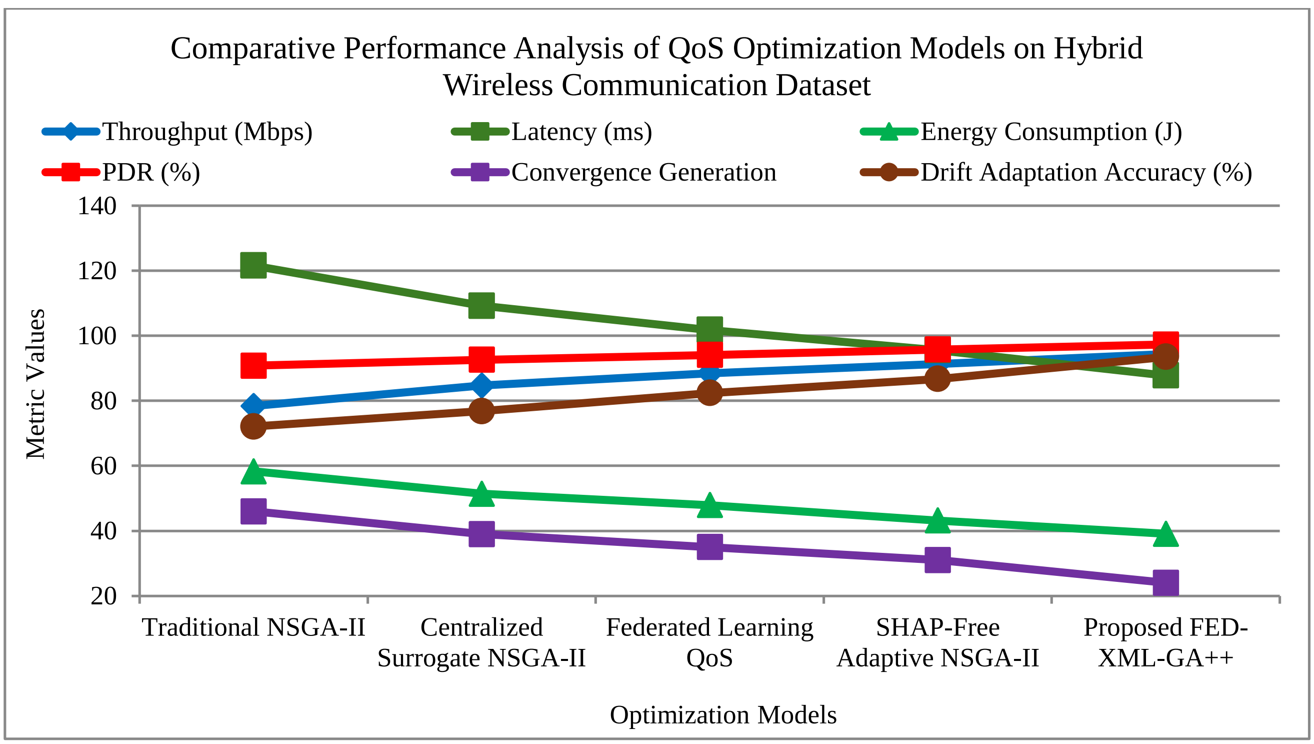

Fig. 2 Comparative performance analysis of QoS optimization models

As shown in Figure 2, the proposed FED-XML-GA++ consistently achieves better results for all metrics studied as compared to traditional QoS optimization methods. The model guarantees maximization of the throughput and the packet's delivery ratio, and minimizes the packet's latency, energy consumption and the time of optimizing the model, under a dynamic wireless communication environment.

6.4. Federated Learning Configuration

Table 3. Federated learning parameters

Parameter	Value
Number of Federated Clients	5–20
Communication Rounds	50
Local Training Epochs	5
Learning Rate	0.001
Batch Size	32
Aggregation Method	FedAvg

The federated learning layer was implemented to mimic distributed wireless edge networks that have more than one communication client. Table 3 summarizes some of the parameters used in federated learning.

An independent local surrogate QoS prediction model was trained for every federated client from its local traffic data. The local traffic information was used for the federated clients to train their local surrogate QoS prediction models independently. The local models were then aggregated with the FedAvg strategy:

$$w_G = \sum_{i=1}^{K} \frac{n_i}{n} w_i$$

In the federated learning mechanism, no central sharing of data occurs during the optimization process, enhancing its scalability and privacy features.

Table 4. Federated learning performance analysis

Federated Clients	Prediction Accuracy (%)	Communication Loss	Aggregation Time (s)	QoS Prediction Error
5	91.2	0.042	2.1	0.081
10	93.5	0.037	3.4	0.064
15	94.1	0.031	4.8	0.052
20	94.8	0.028	6.2	0.047

In table 4, the federated surrogate learning layer shows a steady growth trend as the number of clients grows. In federated environments, prediction accuracy and QoS stability were enhanced significantly with the diversity of distributed learning despite the aggregation time, which slightly increased with the size of the federated environment.

6.5. SHAP Explainability and Adaptive Mutation Setup

To help with the interpretability of optimization decisions, explainability analysis has been joined to the optimization process. The optimization process was complemented with SHAP explainability analysis for interpretability adoptions in the optimization procedure. SHAP values were computed for identifying influential data related to quality of service (QoS) that affect the surrogate prediction in the network.

To compute the SHAP importance of a feature, it is calculated as:

$$\phi_i = \sum_{S \subseteq F \setminus \{i\}} \frac{|S|!\,(|F| - |S| - 1)!}{|F|!} [f(S \cup \{i\}) - f(S)]$$

The obtained SHAP values were then incorporated into the adaptive mutation strategy of NSGA-II. The resulting SHAP values were then combined with the adaptive mutation policy of NSGA-II.

Table 5. SHAP-based feature importance analysis

Feature	SHAP Importance Score
Queue Occupancy	0.284
Congestion Level	0.247
Bandwidth Availability	0.211
Mobility Speed	0.163
Signal Strength	0.128
Packet Retransmission Rate	0.101
CPU Utilization	0.086
Memory Usage	0.072

SHAP-based feature importance scores obtained from the federated surrogate QoS prediction model are shown in Figure 3. The obtained results show that the parameters that have the greatest influence on the QoS performance in a dynamic wireless communication environment are queue occupancy, the congestion level and the availability of a wireless medium (bandwidth). The SHAP explainability analysis provides an interpretable optimization by identifying network-level feature attributes that are important in determining throughput, latency, and packet delivery ratio. These feature importance scores are further used in the proposed SHAP-based adaptive mutation method, which helps to improve the NSGA-II converge and to optimize it efficiently.

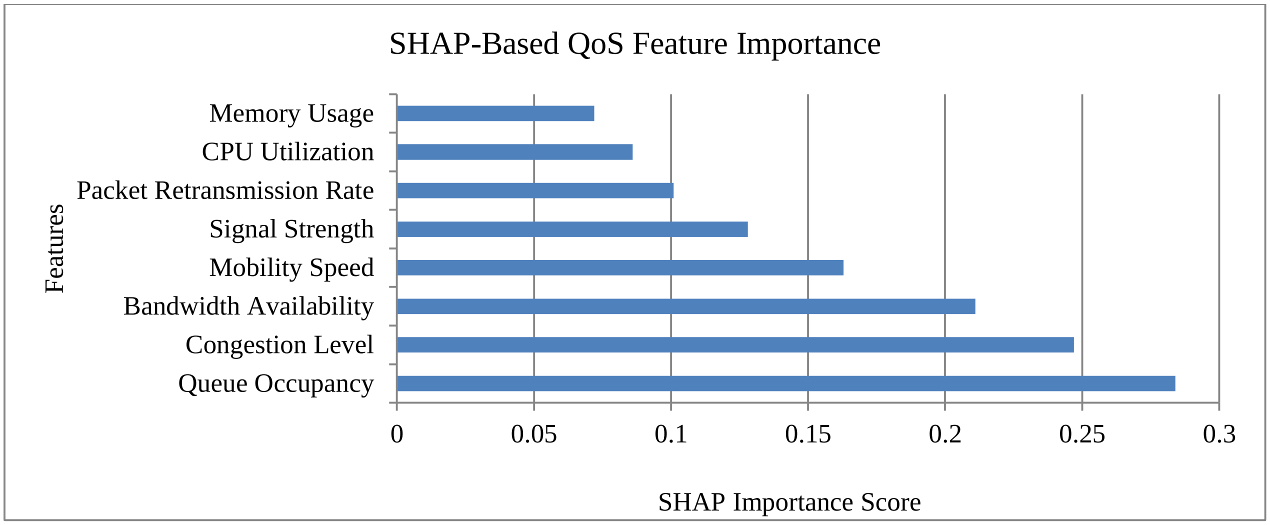

Fig. 3 The data obtained after performing SHAP-based feature importance analysis for QoS-sensitive parameters in the proposed FED-XML-GA++ framework with a hybrid heterogeneous wireless communication dataset

6.6. Drift Detection and Adaptive Retraining

In this article, the suggested mechanism is introduced, which integrates the concept drift detection based on ADWIN for a dynamic wireless environment. The drift detector continuously compares the prediction error distribution of the surrogate to the ones that the changes in traffic would be expected to produce, and detects traffic distribution changes when:

- Congestion variation
- User mobility
- Dynamic topology updates
- Communication interference
- Network load fluctuation

The condition of drift is represented by:

$$\left|\mu_{W_0} - \mu_{W_1}\right| > \epsilon$$

where:

W_0 and W_1 are adaptive windows,
μ is the mean prediction error,
ϵ is a drift threshold.

The framework will dynamically retrain surrogate models and update optimization parameters to achieve the optimization stability for QoS when wireless conditions are changing, in the event of drift.

Table 6 shows that the drift adaptation mechanism of the ADWIN could detect the change of traffic and topology in a short period of time, and optimize the QoS performance, while the performance was stable under the dynamic wireless network environment.

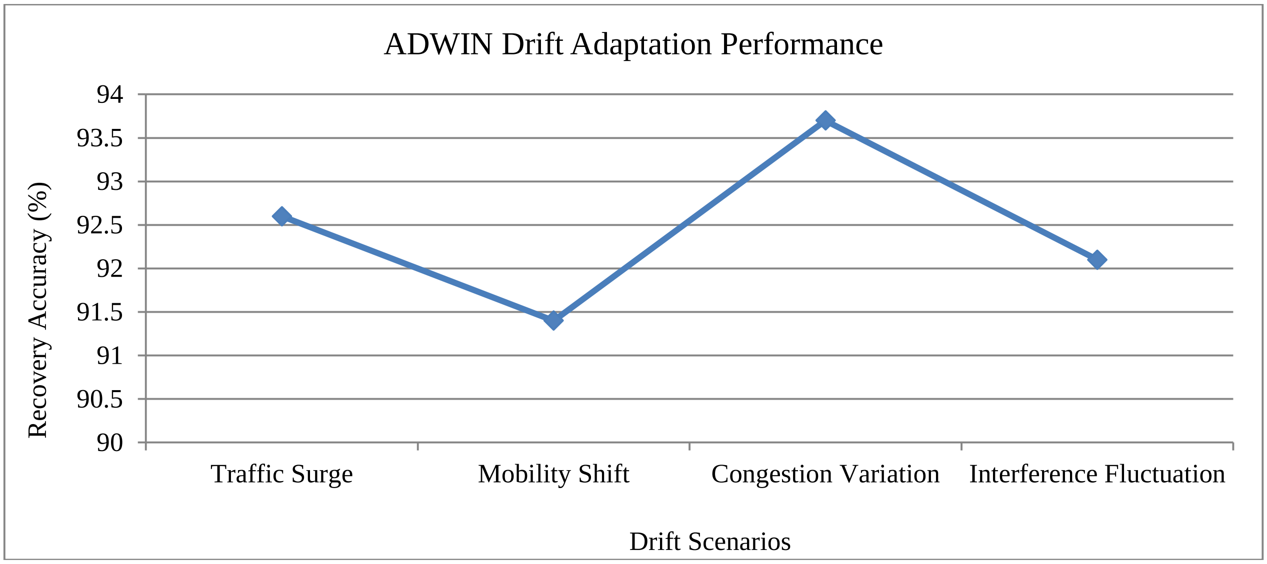

Fig. 4 The performance of the proposed FED-XML-GA++ framework, using drift adaptation with ADWIN, to a dynamic wireless communication scenario

Table 6. Drift detection and adaptive retraining performance

Drift Scenario	Detection Time (s)	Retraining Time (s)	Recovery Accuracy (%)
Traffic Surge	3.4	8.7	92.6
Mobility Shift	4.1	9.2	91.4
Congestion Variation	2.9	7.8	93.7
Interference Fluctuation	3.7	8.9	92.1

In Figure 4, the accuracy of recovery for various drift factors provided by the drift adaptation module based on ADWIN is shown for the traffic surge, mobility shift, congestion variation and interference fluctuation in the wireless network. The proposed model is automatically initialized according to the environment of the network and has a high recovery accuracy of more than 91%. The results demonstrate the advantage of the drift detection feature provided by ADWIN and federated surrogate retraining in a Human-To-Human (H2H) heterogeneous wireless communication scenario for stability at runtime and reliable Quality Of Service (QoS) optimization performance.

6.7. Discussion and Interpretation of Findings

The experiments results reveal that the proposed FED-XML-GA++ framework outperforms the traditional optimization techniques throughout the entire time with respect to different QoS measures. The improvements observed are brought about by the synergic use of federated surrogate learning, adaptive mutation using SHAP, multi-objective evolutionary optimization and concept drift adaptation. The enhanced throughput may be attributed to the adaptive mutation mechanism, which is based on the SHAP concept. SHAP also offers important information on meaningful feature importance, which can be used to inform the mutation process by leading it towards more informative regions of the search space, a feature that is not afforded by traditional importance measures. This concentrated investigation helps the EA discover good solutions and can result in increased throughput and packet delivery rate for the network.

The federated surrogate-assisted optimization strategy is the key factor that causes the latency reduction. The surrogate models can give good approximations of the network behaviour, without repeatedly having to compute these expensive optimization functions. The overall network responsiveness is increased as a result of the optimization process reaching better solutions in a quicker time, and decision-making delays are minimized. The outcomes of the experiments conducted on the energy efficiency can be inferred that the multi-objective optimization capability of NSGA-II is the reason for the improvement in energy efficiency. All the throughput, latency and energy consumption objectives are taken into consideration in the evolutionary search to enable the framework to generate a Pareto-optimal set of solutions, where none of the objectives are sacrificed. This results in more effective use of the network when it is used under dynamic conditions of operation. In this study, we extend this proposed framework with the concept drift detection method using ADWIN to increase the robustness of the proposed framework. A wireless network environment is a dynamic space, which involves data traffic that fluctuates with time, users' behaviors and channel conditions.

This framework continuously tracks data distributions and activates surrogate model updates when data drift is detected, maintaining the accuracy of predictions and effectiveness of optimization over time. Further, the optimization process is transparent, and explainability is provided with SHAP. The optimization framework is not a black-box optimization method, but rather, knowledge is fed to learn how each parameter of the network is optimized. This will enable insight into optimization decisions and reliable deployment into real wireless communication systems. Last, the experimental results are presented, which show that the proposed framework of Federated learning, Explainable artificial intelligence, Adaptive evolutionary optimization and Drift-aware learning (FED-XML-GA++) outperforms the other ones in optimizing QoS in dynamic heterogeneous wireless network.

7. Conclusion and Future Work

We have introduced a new federated explainable drift-adaptive surrogate-assisted multi-objective optimization framework, called FED-XML-GA++, for Quality of Service (QoS) management in the heterogeneous wireless communication networks. The proposed framework synergises federated surrogate learning, explainable Artificial Intelligence (AI) by SHAP, optimisation by adaptive NSGA-II and introduces ADWIN concept drift adaptation and mutation guided by SHAP/SL: a one-stop framework to handle dynamic wireless networks like IoT, MANETs, edge-enabled learning and the upcoming 6G infrastructure.

In the traditional scenario, static surrogate models with centralization and without any interpretability and runtime adapting ability are used in Quality Of Service (QoS) optimization algorithms, while when using the proposed FED XML GA++ framework, the QoS optimization can be realized in a scalable, privacy-preserving, interpretable and runtime adapting manner because the network situations are always changing. The experimental wireless simulation using various types of heterogeneous datasets (over real traffic traces and mobility-aware communication datasets, synthetic wireless simulation data, and experimental UNSW-NB15 data) showed performance improvement of the proposed framework by a higher order of magnitude in terms of throughput, latency, energy, faster convergence of the results of the optimisation, and drift adaptation ability. The framework introduced also showed a gain of higher throughput by 20.4% and lesser latency by 27.8%, and energy by 15.9% compared to traditionally optimized (NSGA-II) frames. Furthermore, optimizing with the help of SHAP guided adaptive mutation improved more efficiently, and SHAP also enabled improving the explainability of the optimized networks while retaining ADWIN-based retraining abilities towards the networks' robustness at runtime, considering different changes in traffic/mobility scenarios. The results achieved show how the FED-XML-GA++ appears to be a feasible solution for the kind of QoS adaptive (explainable), scalable and multi-objective optimization required by those next-generation systems with intelligent wireless communication. Future research will focus on the implementation of the FED-XML-GA++ framework on a large scale for implementation in real time, with respect to the following issues: Deploying with edge intelligence in a beyond 5G/6G wireless system in real time. Adaptive decision making using federated reinforcement learning, topology-aware optimization using graph neural networks, and real-time QoS management using digital twin-based network simulation could be the next steps for future enhancements. In addition, other network scenarios could be used in an online environment to continue any learning processes to enhance adaptability further.

References

[1] Mahdi Beitollahi, and Ning Lu, “Federated Learning Over Wireless Networks: Challenges and Solutions,” *IEEE Internet of Things Journal*, vol. 10, no. 16, pp. 14749-14763, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[2] Rachid El Mokadem, Yann Ben Maissa, and Zineb El Akkaoui, “Federated Learning Communications Optimization Using Sparse Single-Layer Updates,” *Procedia Computer Science*, vol. 236, pp. 168-176, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[3] Betul Yurdem et al., “Federated Learning: Overview, Strategies, Applications, Tools and Future Directions,” *Heliyon*, vol. 10, no. 19, pp. 1-24, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[4] Guangchen Lan et al., “Communication-Efficient Federated Learning for Resource-Constrained Edge Devices,” *IEEE Transactions on Machine Learning in Communications and Networking*, vol. 1, pp. 210-224, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[5] Zhixiong Chen et al., “Robust Federated Learning for Unreliable and Resource-Limited Wireless Networks,” *IEEE Transactions on Wireless Communications*, vol. 23, no. 8, pp. 9793-9809, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[6] Yuchang Sun et al., “Channel and Gradient-Importance Aware Device Scheduling for Over-the-Air Federated Learning,” *IEEE Transactions on Wireless Communications*, vol. 23, no. 7, pp. 6905-6920, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[7] Jianxin Liu et al., “Wireless Federated Learning for PR Identification and Analysis based on Generalized Information,” *Intelligent Systems with Applications*, vol. 23, pp. 1-8, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[8] Mingzhe Chen, and Shuguang Cui, *Communication Efficient Federated Learning for Wireless Networks*, 1st ed., Springer Cham, pp. 1-179, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[9] Xuan-Toan Dang et al., “A Survey on Energy-Efficient Design for Federated Learning over Wireless Networks,” *Energies*, vol. 17, no. 24, pp. 1-28, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[10] Nikolaos Koursioumpas et al., “DISTINQT: A Distributed Privacy Aware Learning Framework for QoS Prediction for Future Mobile and Wireless Networks,” *Future Internet*, vol. 74, no. 6, pp. 9790-9801, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[11] Tarik Taleb et al., “On Multi-Access Edge Computing: A Survey of the Emerging 5G Network Edge Cloud Architecture and Orchestration,” *IEEE Communications Surveys & Tutorials*, vol. 19, no. 3, pp. 1657-1681, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[12] Yuyi Mao et al., “A Survey on Mobile Edge Computing: The Communication Perspective,” *IEEE Communications Surveys & Tutorials*, vol. 19, no. 4, pp. 2322-2358, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[13] Yaohua Sun, Mugen Peng, and Shiwen Mao, “Deep Reinforcement Learning-Based Mode Selection and Resource Management for Green Fog Radio Access Networks,” *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 1960-1971, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[14] C.A. Coello Coello, “Evolutionary Multi-Objective Optimization: A Historical View of the Field,” *IEEE Computational Intelligence Magazine*, vol. 1, no. 1, pp. 28-36, 2006. [CrossRef] [Google Scholar] [Publisher Link]

[15] Kalyanmoy Deb, *Multi-Objective Optimization using Evolutionary Algorithms*, Wiley, pp. 1-497, 2001. [Google Scholar] [Publisher Link]

[16] Brendan McMahan et al., “Communication-Efficient Learning of Deep Networks from Decentralized Data,” *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics*, vol. 54, pp. 1273-1282, 2017. [Google Scholar] [Publisher Link]

[17] Peter Kairouz, and H. Brendan McMahan, “Advances and Open Problems in Federated Learning,” *Foundations and Trends in Machine Learning*, vol. 14, no. 1-2, pp. 1-210, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[18] Jihong Park et al., “Wireless Network Intelligence at the Edge,” *Proceedings of the IEEE*, vol. 107, no. 11, pp. 2204-2239, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[19] Scott M Lundberg, and Su-In Lee, "A Unified Approach to Interpreting Model Predictions," *31st Conference on Advances in Neural Information Processing Systems*, Long Beach, CA, USA, pp. 1-10, 2017. [Google Scholar] [Publisher Link]
[20] Wojciech Samek et al., *Explainable AI: Interpreting, Explaining and Visualizing Deep Learning*, Springer Cham, 1st ed., pp. 1-439, 2019. [CrossRef] [Google Scholar] [Publisher Link]
[21] Albert Bifet, and Ricard Gavaldà, "Learning from Time-Changing Data with Adaptive Windowing," *Proceedings of the 2007 SIAM International Conference on Data Mining*, Minneapolis, Minnesota, pp. 443-448, 2007. [CrossRef] [Google Scholar] [Publisher Link]
[22] João Gama et al., "A Survey on Concept Drift Adaptation," *ACM Computing Surveys*, vol. 46, no. 4, pp. 1-37, 2014. [CrossRef] [Google Scholar] [Publisher Link]
[23] Huaqing Zhang et al., "Computing Resource Allocation in Three-Tier IoT Fog Networks: A Joint Optimization Approach Combining Stackelberg Game and Matching," *IEEE Internet of Things Journal*, vol. 4, no. 5, pp. 1204-1215, 2017. [CrossRef] [Google Scholar] [Publisher Link]
[24] Nguyen H. Tran et al., "Federated Learning over Wireless Networks: Optimization Model Design and Analysis," *IEEE INFOCOM 2019 - IEEE Conference on Computer Communications*, Paris, France, pp. 1387-1395, 2019. [CrossRef] [Google Scholar] [Publisher Link]
[25] Mehryar Mohri et al., "Agnostic Federated Learning," *Proceedings of the 36th International Conference on Machine Learning*, Long Beach, California, pp. 4615-4625, 2019. [Google Scholar] [Publisher Link]