Journal of Wireless Sensor Networks and IoT
Journal of Wireless Sensor Networks and IoT, ISSN: 3048-8729 Vol. 3, No. 2, 2026 (pp. 1-10)
RESEARCH ARTICLE
ECEJOURNALS.IN

Edge-Intelligent Wireless Sensor Networks: A Federated Learning Framework for Energy-Aware Distributed Inference

Dahlan Abdullah*

Department of Informatics, Faculty of Engineering, Universitas Malikussaleh, Aceh, Indonesia.

KEYWORDS:

Wireless Sensor Networks,
Federated Learning,
Edge Intelligence,
Energy-Aware Optimization,
Distributed Inference,
Network Lifetime Enhancement

Author's Email:
dahlan@unimal.ac.id

DOI: 10.31838/WSNIOT/03.02.01

Received : 10.12.2025

Revised : 05.01.2026

Accepted : 02.02.2026

Abstract

The internet of things (IoT) is also using sensor networks that are increasingly becoming wireless as well as incorporating edge computing to achieve distributed inference instead of processing the data through a centralised cloud computing service. While, the traditional federated learning (FL) models are not specifically optimised to operate within energy constraints characteristic of battery-operated sensor nodes, which results in untimely depletion of nodes and network saturation. This paper: this paper suggests a federated learning model that is energy- aware on edge-intelligent WSNs that optimises computation and communication overhead on joint distribution training. A mathematical model of energy consumption is formulated that models node-based computation and transmission costs and re-formulates the global learning optimization with hard constraints on the available energy. The adaptive mechanism of participation is proposed, where only those nodes that have enough residual energy and good channel conditions are eligible in contributing to every training round. Moreover, update schemes that are communication efficient are incorporated to minimise further the transmission overhead. Large-scale simulations show that the suggested framework might save as much as 28 % of the total energy usage as compared to typical FL with equivalent inference competence. The methodology also increases the network life by more than 35 % and convergence speed in non-homogeneous node environments. The above findings suggest that sustainable edge intelligence in costly wireless sensor networks needs an energy-conscious federated optimization.

How to cite this article: Abdullah D (2026). Edge-Intelligent Wireless Sensor Networks: A Federated Learning Framework for Energy-Aware Distributed Inference. Journal of Wireless Sensor Networks and IoT, Vol. 3, No. 2, 2026, 1-10

Introduction

WSNs has changed the traditional infrastructures that were invariably passive sensory to be considered as intelligent distributed systems that can undertake localised data processing and inference in real-time [4], [8]. Combining edge intelligence enables sensor nodes and gateway to perform machine learning operations close to the source of data, cutting down on latency, and decreasing bandwidth requirements and making autonomous decisions and normal decisions.[1, 8] These features are becoming imperative in the use of environmental monitoring, precision agriculture, industrial process control and smart infrastructure management. Nonetheless, implementing the learning mechanisms in WSNs poses a great challenge because sensor nodes have very low resource constraints. The most limiting factor in large scale WSN deployments has been energy availability. The battery capacity of most sensor nodes is limited or they use intermittent energy sources such as solar, wind and thermal energy, thus making sustainable operation a major design goal. In a distributed learning environment, repetitive local training and exchange of model parameters through iteration substantially raises the cost of computations and a exchange of model parameters. A significant part of the overall energy consumption, in particular, is covered by wireless transmission. Unless properly optimised, distributed inference well will speed-up node depletion, sensing coverage and overall network lifetime will still be shortened. Federated learning (FL) has become an encouraging concept of distributed model training without any centralised data aggregation.[9] FL helps decrease the availability of raw information, as one or two of them end up in just a single place on it, and small pieces of the model can be shared in their place, thereby lowering privacy risks and eliminating the overload of data transmission.[3, 5] However, traditional FL algorithms are most likely to be trained on mobile-based applications or edge-cloud systems where the issue of energy is relatively medium.[12] These models tend to presuppose equal participation put by clients, synchronous communication cycles, and unrestricted iterative training models.[2, 6] These assumptions contribute to poor use of energy, irregular levels of node usage and poor performance of the network in the long-term when applied directly to battery-powered WSN settings. Even though several previous studies have already analysed energy-efficient communication protocols in WSNs and communication-compression techniques in federated learning,[6, 7] there is little literature that explicitly analyses the federated optimization mechanism as requiring explicit energy modelling. Current methods usually optimise both learning performance and network energy consumption separately and not as closely coupled design criteria.[7, 11] Such a disconnection forms a research gap toward effective provision of sustainable edge intelligence functions within feasible sensor network conditions. To handle this shortcoming, this paper proposes an energy conscious federated machine learning architecture, which is specifically optimised towards the edge-based WSNs. An exhaustive node model energy consumption is developed to collectively define both the computing and transmission costs when federated training is involved. Reformulation of global optimization objective is done to include explicit energy limits guaranteeing high energy participation among the nodes in the communication rounds in a sustainable and balanced way. On this formulation, a mechanism of adaptive client selection is proposed that uses the residual energised states and communication environments in dynamically controlling participation, hence balancing inference accuracy and network persistence.[10, 12] The proposed framework is also justified using scalable simulations involving better convergence rates, drastic decreases in total energy usage, and drastic duration of network lifetime compared to the traditional federated learning schemes. The proposed scheme provides a sustainable channel towards the implementation of intelligent analytics in large-scale wireless sensor network by closely integrating distributed learning optimization with network design taking into consideration energy consumption.

Related Work

Federated learning has become a popular distributed learning paradigm to be applied in Internet of Things (IoT)-based and edge computing settings.[1, 8] FL ensures that privacy risks are minimised through collaborative optimization of models, without communication smoothness where data is aggregated in a central location.[3, 9] More recent work has discussed lightweight federated design to edge and asynchronous aggregation to reduce straggler effects and to have communication efficient update mechanisms including gradient sparsification and quantization.[2, 6, 7] FL is used in IoT-based systems in the contexts of smart healthcare monitoring, fault diagnosis in industries, and sensorimotor environments.[12] Although these solutions have been made, the vast majority of federated system models are developed on the basis of relatively constant computational and energy capabilities, which frequently fails to reflect the hard battery constraints of wireless sensor networks.[5, 11] Simultaneously with the advances in distributed learning, a large amount of research has been performed to develop energy-efficient WSNs procedures. Classical strategies are aimed at reducing the radio transmission costs by using clustering mechanisms, duty-cycling schemes, adaptive power control of transmission, and energy-conscious routing methods. Certain protocols like LEACH and its derivatives add hierarchical communication in order to balance the load between the nodes, whereas cross-layer optimization approaches seek to coordinate between medium access and routing choices. Though they are viable to minimise communication overhead, these techniques only emphasise on sensing and data forwarding capabilities, but not on iterative learning workloads, which also add new computational and synchronisation requirements.

Unexpectedly, the inference mechanisms to be distributed have been also explored to minimise dependency on centralised processing in sensor networks.[4, 8] Strategies are in-network aggregation, consensus based estimation, distributed detection and collaborative filtering model. More recently, edge-assisted inference models have been used to execute local models, which run on the gateway nodes, or cluster heads to minimise the latency and backhaul traffic.[1, 12] Nevertheless, most of these mechanisms are concerned with the inference once the training of the model is over, without considering the energy implications of distributed training itself. Thus, the training step, and especially, in the case of federated optimization, is not optimised sufficiently in terms of energy sustainability.[6, 7] A number of recent papers are trying to integrate energy-awareness into federated learning, which includes client selection heuristics, communication compression, or responsive training frequency.[10, 12] Although these techniques partially decrease overhead, they may be based on simplified models related to the overhead like constant energy levels, equal capacities of nodes, or failure of detailed computation-communication interactions.[7, 11] Additionally, energy optimization is often considered as a secondary constraint other than an objective part of global learning.[2, 5] Such discontinuity in treatment does not enable the scalability of federated learning and its long-term viability in densely deployed WSN settings. In general, the current body of literature shows significant advancement in federated optimization, energy-efficient networking on a case-by-case basis, but an overview system that strictly combines node-level energy modelling, constrained federated optimization, and adaptive participation control has not been studied yet. This lack of an such integrated approach denotes a faulty gap in facilitating sustainable edge wisdom to battery-powered remote wireless sensor networks. This void is filled with the current work creating a mathematically justified energy-conscious federated learning model, which explicitly certifies the distributed inference performance with the optimization of network lifespan.

Methodology

System Model and Energy Formulation

The proposed wireless sensor network model, based on edge-intelligence is shown in Figure 1, which is a hierarchical three-layer structure that comprised of distributed sensor nodes, a cluster head (gateway) and an edge server that is the global aggregator.

Sensor nodes at the most basic level (i.e., the bottom) store local datasets , local model parameters , and finite residual energy . The nodes also do a local training on their local dataset and send the model updates (not the raw data) on a periodic basis. The middle layer will consist of a cluster head which will organise the communication as well as implement an energy conscious participation policy. In every round of communication, the gateway cheques the residual

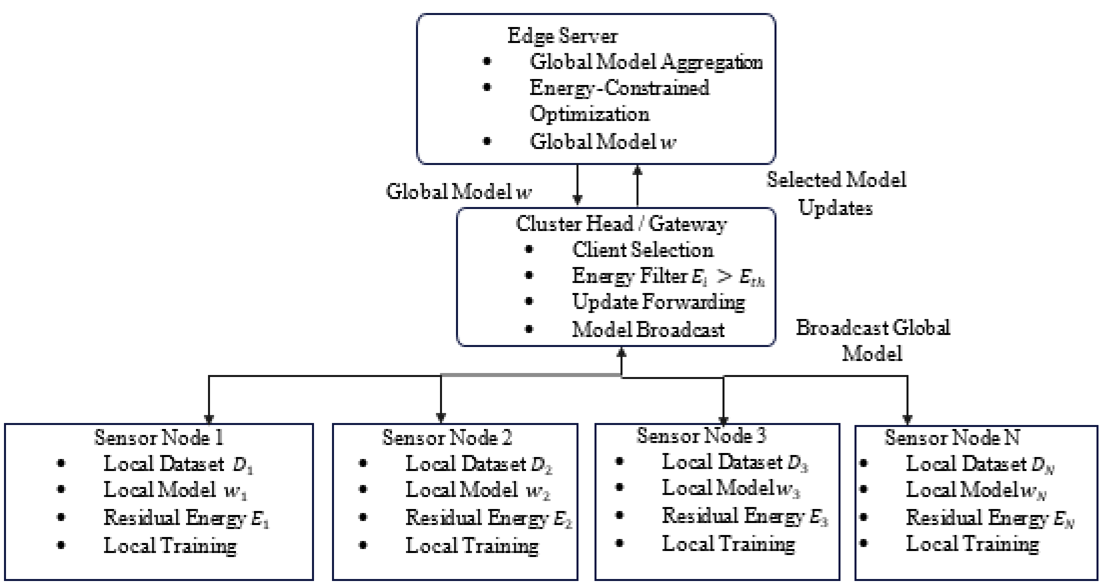

Figure 1: Edge-Intelligent WSN Architecture

energy of the communicating nodes and only allows training under the condition . Only selected updates are sent to the edge server where global aggregation is done. The global model is then communicated to capable nodes, in a form of aggregation bringing the communication of federated learning back to its source. As Figure 1 illustrates, the energy-conscious client filter is directly incorporated into the update stream and allows one to sustainably perform distributed inference with limited power budgets. In order to be able to quantitatively model the resource consumption, the total energy used by node in one round of federated learning is modelled as the aggregate of components of computation, transmission and reception energy consumptions. Total energy consumption is given as

$$E_i = E_i^{comp} + E_i^{tx} + E_i^{rx} \quad (1)$$

Where E_i^{comp}, E_i^{tx}, and E_i^{rx} denote computation, transmission and reception energy respectively. Local training would consume a amount of computation energy and is dependent on the features of a processor and workload complexity and is modelled as.

$$E_i^{comp} = \kappa C_i f_i^2$$

in which, κ is the effective switched capacitance coefficient, C_i is the number of CPU cycles to do local model updates, and f_i is the operating frequency of node i. The energy spent in uploading local model changes to the gateway is as shown by

$$E_i^{tx} = P_{tx} \cdot \frac{S_i}{R_i}$$

In which represents the transmission power, represents the size of the implemented model parameters, and represents the available data rate. The combination of these formulations has the benefit of modelling the interaction between computation and communication overhead during federated learning iterations, which is realistic given operational limitations in wireless sensor networks. In global federated learning, the challenge of reducing the cumulative loss among distributed nodes without disproportionately compensating any individual is the global federated learning objective that needs to be reduced. The opposite end is optimised as follows:

$$\min_w F(w) = \sum_{i=1}^{N} \frac{n_i}{n} F_i(w) \quad (2)$$

refers to the local loss of node , is the local samples, is the total number of nodes, which is optimised by computing: In order to ensure sustainable operation of the network, the goal in Equation (2) is redefined with clear energy limitation. The energy optimization problem is outlined as an energy-conscious problem.

$$\min_w F(w) \; s.t \;\; E_i > E_{th} \quad (3)$$

is the specified minimum energy necessary to join. The nodes with energy smaller than this value are not involved in the current training round and this will allow network connectivity to be maintained by avoiding early depletion. The designed framework makes the provision of the constraint as part of the federated learning process by providing a principled balance between the accuracy of distributed inference and long-term network sustainability.

Federated Learning Framework Proposal.

This paper makes a suggestion of an Energy-Aware Federated Learning (EA-FL) framework that aims to maintain distributed inference within WSNs, shared by controlling participation and communication overhead. The eligibility score assigned to a node at the start of a round is the amount of residual energy available at node iii, calculated by local battery reading or energy harvesting state, and allows the node to participate in completing local training and transmitting updates without exceeding the constraint on the amount of energy that node can consume in a given round as indicated in Equation (3). Along with energy, the network also includes channel-aware scheduling that takes the link quality indicators (e.g., SNR, RSSI, or achievable rate when reducing retransmissions and eliminating the use of nodes with bad uplink conditions which can generate excessive transmission energy. The general formulation of EA-FL, on these two factors, is a probability of participation model which tend to choose nodes of larger residual energy, and better channel condition, whereas, low-energy, or weak-link nodes are deferred or disqualified to avoid premature network fragmentation. This probabilistic sampling is used in order to prevent repeated sampling of the similar superior nodes, and thus, balance learning progress and equity and lifetime. Generally, to minimise further the communication energy, EA-FL incorporates mechanisms of communication reduction into the update pipeline. To begin with, gradient sparsification benefits by only transmitting top- most important elements of the gradient or model deltas, which drastically reduces the amount of payload in the transmission energy model. Second, quantization codes updates sent by compressing them (e.g. 8-bit or stochastic quantization) but maintaining convergence stability, and again save on communication cost. Third, periodical aggregation also minimises the number of uplink transmissions per local epoch by enabling the aggregation of many local epochs beforehand or by aggregating only after every T rounds, thereby reducing the overall number of communication rounds and achieving additional lifetime enhancement. These processes are implemented either in isolation or together based on constraints of a system in a way that energy savings are not attained at an unacceptable cost to inferences. The overall training process can be summed up with the help of the Algorithm 1 due to the description of the EA-FL model of energy assessment, adaptive customer selection, and compressed transmission of local updates, weighted global aggregate, and update in the residual energy.

Algorithm 1: Energy-Aware Federated Learning (EA-FL)

Input: Number of rounds *T*, node set *N*, energy threshold E_{th}, local epochs *E*, learning rate *η*, compression settings (sparsity *k*, quantization bits b), aggregation period *τ*

Output: Trained global model w_T

1. **Initialize** global T, model at edge server
2. **for** round to **do** w_0
3. Edge server **broadcasts** current global model w_{t-1}t to gateway and nodes
4. **Energy evaluation:** each node $i \in N$ estimates residual energy and link quality (e.g., rate
5. **Client filtering:** gateway forms eligible set $S_t=\{i \in N \mid E_i > E^{th}\}$
6. **Channel-aware scheduling:** compute a selection score $q_i(t)$ using energy and channel quality
7. **Participation sampling:** select $K_t \subseteq S_t$ according to $q_i(t)$ (probabilistic or top-mmm)
8. **Local update:** each selected node $i \in K_t$ performs local training for E epochs $w_i^t \leftarrow \eta \nabla F_i(w_{t-1})$
9. **Compression:** each node applies sparsification + quantization to the update $\Delta_i^t = w_i^t - w_{t-1}$
10. Selected nodes **upload** compressed updates $\hat{\Delta}_i^t$ to gateway (and gateway forwards to edge server)
11. **Periodic aggregation:** **if** τ = 0 then
12. Edge server performs **weighted aggregation** $w_t \leftarrow w_{t-1} + \sum_{i \in \mathcal{K}_t} \frac{n_i}{\sum_{j \in \mathcal{K}_t} n_j} \hat{\Delta}_i^t$
13. **else** set $w_t \leftarrow w_{t-1}$
14. **Energy update:** each node updates residual energy $E_i(t+1) \leftarrow E_i(t) - \left(\boldsymbol{E}_i^{comp} + \boldsymbol{E}_i^{tx} + \boldsymbol{E}_i^{rx}\right)$
15. end for
16. **Return**

Complexity Scalability Analysis.

The communication and computational expenses of the suggested Energy-Aware Federated Learning (EA-FL) architecture also have a direct effect on the applicability of the framework to large-scale application in wireless sensor networks. The given part of the research examines the complexity aspect and scalability of the framework when the node density is growing and the energy limitations are different. The complexity of the communication per federated learning round mainly relies on the number of participating nodes and the size of model updates that are being presented. Let denote the total amount of nodes and S denote the mean size of model parameters or gradients sent. In the extreme situation of every node involving itself in a particular round, the uplink communication complexity can be defined as

$$O(N \cdot S)$$

Nevertheless, using the suggested energy-conscious client selection mechanism, only a subgroup $K_t \subseteq N$ takes part in each round, literally lowering the essential communicate cost $O(|K_t| \cdot S)$, thus diminishing the prevailing communication energy aspect, which is expressed in Equation (1). In its turn, the proposed framework displays sub-linear practical communication development compared to the total node count under the condition of participation filtering. The complexity of computation is the result of the number of local training epochs and batch operations performed on each node. Where E is the number of local epochs and B is the mean size of the batch or data processing factor of each local epoch. The general computational work per round of all the participating nodes is approximately.

$$O(N \cdot E \cdot B)$$

Just as is the case with communication cost, effective computational burden is depowered to $O(|K_t| \cdot E \cdot B)$when there is participation restriction using energy. This sort of scaling reduces the energy depletion that is uniform and allocates dynamically the training weight in the network. There is a basic compromise between energy conservation and model accuracy. Convergence acceleration of increasing local epochs E would minimise the number of communication rounds and but consequently consume more computation energy E_i^{comp}. In contrast to this, aggressive participation filtering will reduce aggregate energy consumption but can effectively reduce convergence as results of smaller model diversity in aggregation. The offered EA-FL framework stabilises this trade-off and provides the outcomes of the dynamical regulation of the participation likelihood and the frequency of aggregation depending on the conditions of residual energy. Another very important consideration is scalability to higher node density. In the traditional federated learning system, the number of nodes is added at a cost in terms of communication overhead and synchronisation complexity. Instead, the proposed energy-worthy participation framework will guarantee that just a limited fraction of nodes will participate in a specific round to avoid congestion of uplink, and preserves a stable convergence pattern. The adaptive client selection which defines how packet switching is allowed reduces the number of active packets as network size increases, and as such, causes the framework to be able to grow to dense WSN implementations without a linear increase in power consumption. E_th is a crucial parameter of network behaviour that can be used in controlling the network. Stepping the threshold to a higher point results in a more rigorous participation requirement, which maintains a long-term existence of the network but could be slowing learning? Reduced threshold enhances the speed of training at short-term duration but is prone to fast depletion of energy and uneven node fatigue. Thus, threshold tuning allows system designers to choose a compromise between data convergence speed and operation life depending on the needs of the application. Through theory, it can be estimated that network lifetime will be the trainings rounds that an essential portion of the nodes will reduce and hit the energy threshold. Since the suggested architectural design selectively does not incur low-energy nodes and allows compressing the size of updates, the amount of decrease in energy per round has been reduced in comparison to traditional federated learning. This regulated energy usage translates to a longer operational life and a better survivability to node failures especially in nonhomogeneous energy settings. Generally, the analysis of complexity shows that the version of the EA-FL framework proposed is easily scalable to the network size while having a principled trade-off between distributed inference accuracy and reliance on sustainable energy consumption.

Experimental Setup

The suggested Energy-Aware Federated Learning (EA-FL) scheme is tested in terms of complex simulations aimed at representing the real-life situation of a wireless sensor network. The network topology will be a random distribution of sensor nodes distributed on a rectangular area of a fixed extent on a square. The quantity of nodes will be 50 to 200 to evaluate the scalability with an ever-thickening network. It presupposes the hierarchical topology, which requires sensor nodes to be connected to a cluster head or a gateway that transmits selected updates to an edge server that will divide them throughout the whole world. Starting values of the heterogeneous energy levels of all nodes are set in order to replicate realistic deployment conditions. In distributed training evaluation, a evaluated designated conduction classification dataset is divided to nodes to emulate non-independent and non-identically distributed (non-IID) local data situations. The nodes form a local subset D_i of the whole dataset, and data heterogeneity is produced by giving the nodes biassed distributions of the classes. This arrangement corresponds to realistic conditions of sensing in which observations are not identical at spatially distributed sensors. The learning model used is a lightweight neural network that could be executed to run on edges and the hyperparameters applied in all the comparative methods are equal to maintain fairness. A distance-dependent path loss formulation in the wireless linkage of sensor nodes and the gateway are modelled with additive white Gaussian noise. The attainable transmission rate for each node is calculated around the signal-noise ratio(SNR), and the channel conditions of different nodes are dissimilar to approach heterogeneous communication situations. The channel modelling directly affects the use of transmission energy as in the definition of Equation (1), in which the communication overhead can be realistically assessed. Parameters used to set the energy are picked to represent the typical sensor hardware; low-power hardware. The preliminary energy of the battery is allocated within a set club to represent variability in the devices. The transmission power , processor frequency , switched capacitance coefficient , and update size Si are set to more specifications of the typical embedded devices. The energy model created in Section 3.1 is used to compute the energy consumed in a single training round by and multi-way Faraday machine, computation energy and communication energy. To control the participation of the nodes within the proposed framework, an energy threshold is established. In order to test the usefulness of EA-FL, there is a baseline carried out in two methods. The first baseline is centralised learning, where all the local data are sent to a central server and trained, which is communication-intensive and model training is not distributed. The second baseline involves the normal federated learning that does not use energy conscious client selection or communication compression. The methods are tested using the same network conditions and training parameters, and the performance should be measured by the results in terms of recognition accuracy, total power consumption, convergence time, and network life in terms of communication rounds. This experimental framework allows conducting a thorough evaluation of the learning performance as well as energy sustainability in terms of network size.

Performance Evaluation and Results.

This part discusses and assesses the suggested Energy-Aware Federated Learning (EA-FL) framework with regards to convergence performance, energy efficiency, scalability, and network lifetime. Table 1 summarised simulation and system configuration in which all experiments were done. They are compared with centralised learning and the conventional federated learning in the same training environment to make it fair. The simulation parameters in Table 1 indicate realistic parameters of a wireless sensor network, such as heterogeneous deployment of nodes, transmission power limitation, processor specification, and an explicit participation energy cut-off threshold. The network size will range between 50 nodes and 200 nodes to cheque the scalability and 100 initial nodes to analyse lifetime to ensure consistency across the performance comparisons.

Figure 2 gives the convergence behaviour of the three approaches and the figure plots the test accuracy with respect to the number of the communication rounds. As seen in Figure 2, centralised learning has the quickest convergence rate and the greatest final accuracy since there is complete data aggregation without distributed constraints. Ordinary federated learning has slightly slow convergence yet consistent performance round to round. There is a slightly slower initial convergence rate of the proposed EA-FL framework because of an energy-wise client filtering, though; it reaches a similar final accuracy in the end.

The last accuracy values provided in Table 2 affirm this assertion because EA-FL reported 90.3% accuracy as opposed to the 90.5% accuracy reported by standard federated learning and 92.4% accuracy reported by centralised learning. The reduced performance drop points to the fact that energy consciousness behaviour does not adversely affect the quality of inferences.

Figure 3 shows energy scalability with node density where total energy consumption is a graph of sensor node versus the many sensor nodes. The highest rate of growth in energy in centralised learning as indicated in Figure 3 is because of communication

Table 1: Simulation and System Parameters

Parameter	Value	Description
Number of Sensor Nodes	50–200	Scalability analysis
Initial Nodes (Lifetime Test)	100	For Figure 4
Deployment Area	500 m × 500 m	Random node placement
Local Epochs (E)	5	Per round
Batch Size (B)	32	Local training
Learning Rate	0.01	SGD optimizer
Transmission Power (Ptx)	50 mW	Radio model
CPU Frequency (fi)	1 GHz	Node processor
Energy Threshold (Eth)	0.4 J	Participation constraint
Aggregation Period (τ)	1 round	Unless stated
Update Compression	Top-k sparsification + 8-bit quantization	EA-FL only

Table 2: Overall Performance Comparison

Method	Final Accuracy (%)	Total Energy @100 Nodes (J)	Network Lifetime (Rounds)
Centralized	92.4	5.8	78
Standard FL	90.5	4.1	92
Proposed EA-FL	90.3	2.9	>100

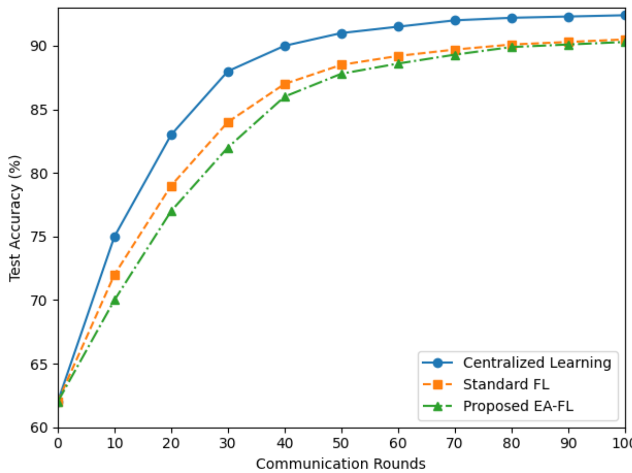

Fig.2: Convergence Performance: Test Accuracy versus Communication Rounds

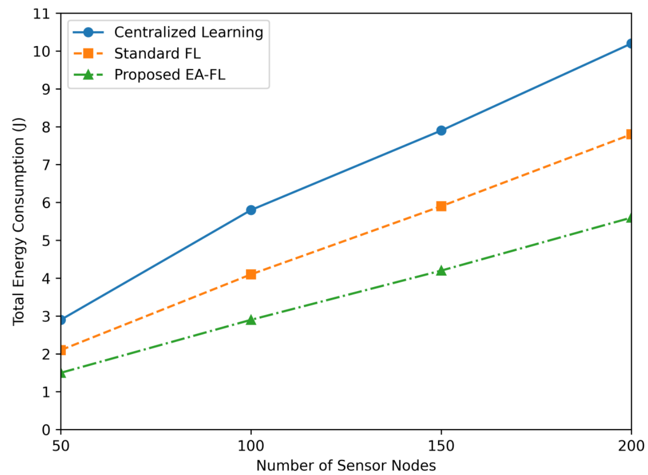

Fig. 3: Scalability Analysis: Total Energy Consumption versus Number of Sensor Nodes

overheads. Standard federated learning saves on the total amount of energy used since there is no transmission of raw data but the energy cost of the process is still proportional to the number of nodes since every node is involved in every round. On the contrary, the offered EA-FL model can save a great amount of energy in every network with a different number of clients, with the help of adaptive selection of clients and compression of communications. Table 2 presents the energy consumption at 100 nodes as consistent with Figure 3 (EA-FL uses 2.9 J of energy versus 4.1 J of energy used in the standard federated learning and 5.8 J of energy used in the centralised learning). Competence of the suggested computation-communication coupling plan that was presented in Section 3 is justified by this decrease.

The applied effect of less energy usage can also be seen in terms of network lifetime, which was analysed in the Figure 4, where the number of active nodes was plotted against the communication rounds. When centralised learning is applied as shown in Figure 4; the node will be emptied very fast because of large transmission overhead thus collapsing the network at an early stage. Normal federated learning enhances lifetime with a moderate degree of efficiency by removing the centralised transfer of data, yet, it suffers progressive node fatigue. The suggested EA-FL model has much more at-work nodes during the training, which prolongs the operational time to over 100 communication rounds. This trend is supported by the lifetime values, as summarised in Table 2 and it is clear that the network functionality of EA-FL continues past 100 rounds, compared to centralised learning that fails past 78 rounds, and federated learning (with standard) that fails past 92 rounds.

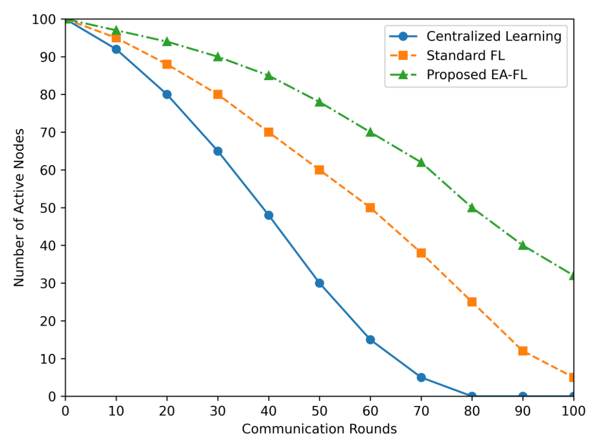

Fig. 4:Network Lifetime Evaluation: Active Nodes versus Communication Rounds

In general, the findings in Figures 2-4 and Tables 1-2 all prove the argument that the proposed EA-FL framework leads to the balanced trade-off in the context of enhancing learning performance and energy sustainability. This framework keeps the accuracy of the competitive convergence constant, but minimises overall energy usage and network lifetime, which makes its use appropriate in far-scale edge-intelligent wireless sensor networks.

Discussion

It can be seen in the results of the experiment that there is an inherent trade-off in energy efficiency versus model accuracy in federated learning under energy constraints. Figure 2 demonstrates that the suggested EA-FL scheme has a slightly slower convergence at the first stages than a conventional federated learning because of the selective participation of nodes and compression of communications. Nonetheless, the ultimate accuracy difference is a minor one, which means that there is no significant inference deterioration caused by the energy-sensitive filtering. This is due to the trade-off that limiting participation will mean that there will be less diversity in the models during the initial rounds, but the same limit will ensure that it does not waste any energy. Their findings indicate that selective energy thresholds have the potential to maintain the stability of long-term learning, and reduce the total amount of energy used as it has been observed in figure 3. The other consideration will be robustness under non-homogeneous node conditions. Practical wireless sensor networks have nodes which vary in residual energy, processing power and quality of channel. The adaptive mechanism of participation delivered in EA-FL, selectively chooses the nodes dynamically in terms of available energy and communication conditions, and hence avoiding uniform energy depletion. In contrast to the conventional federated learning, where in some cases, the selection of high-capability nodes can occur repeatedly, making it faster, the framework proposed makes the workload of the network more balanced. This fact is indicated by the depleting curve of nodes in Figure 4 using a gradual node depletion pattern whereby EA-FL maintains more active nodes as well as a larger number of communication rounds. Resilience of the network is boosted by the capacity to embrace heterogeneity and minimises the chances of untimely fragmentation. With regards to deployment, the framework can be used with hierarchical WSN architecture, which already uses cluster heads or gateways. The extra computational cost of computing power in energy monitoring and client condition is very low by comparison with local model training costs. Besides, gradient sparsification and quantization combined enable the minimization of transmission payload size without hardware specifics. The suggested method is viable because these attributes allow it to be implemented on resource-limited embedded systems that are frequently utilised by IoT and environmental sensing purposes. Nevertheless, meticulous parameter consideration, such as the energy threshold and aggregation rate is a balancing factor between convergence rate and network lifetime in practise. As much as it has its strength, it has a number of limitations. First, the present analysis takes place in a simulation setting and it lacks the real-life elements of accidental interference, hardware malfunction or variable variable of dynamic energy harvesting. Second, the energy model is based on simplified processor and transmission characteristics which might be differing between hardware platforms. Third, although compression of communication is efficient, it may lead to convergence instability of very non-IID data distributions that are overly sparsified. Lastly, there is no reference to the security and adversarial robustness issues and they will have to be examined in upcoming research. On the whole, as it has been pointed out, a way out of the need to incur high energy usage is to directly implement energy-aware systems in federated optimization so as to offer a viable solution in the context of sustainable edge intelligence in the wireless sensor network, and in fact illustrate a way ahead to further implementation and experimentation.

Conclusion

The present paper outlined an energy-conscious federated learning system suitable to federated edge-intelligent wireless sensor networks with severe power consumption requirements. A hierarchical system architecture that merged sensor nodes, gateway-based client selection, and edge-level aggregation was created so as to realise sustainable distributed inference. An extensive node average power model representing the computation and communication cost was designed and the federated optimization problem modified to have explicit energy constraints. Out of this formulation, adaptive client selection strategy along with reduces the communication had been added to strike a balance between model accuracy and long-term viability of the network. The experimental findings showed that the framework presented attains the performance of a competitive convergence with a major impact on a decrease in energy consumption. The proposed approach achieved about 29 % of energy savings at 100 nodes in comparison to standard federated learning with almost the same final accuracy. Also, analysis of network lifetime revealed that the energy-conscious strategy allows the network to operate longer than 100 communication rounds, compared to centralised learning as well as more traditional federated strategies. These findings affirm the fact that it is possible to enforce the concept of energy-awareness directly into the optimization loop of the federation based on distributed learning in order to achieve viable trade-off associated with the distributed learning performance and the sustainable functioning of the network. Future research agenda is to incorporate secure federated learning machinery to counter the adversarial attacks and privacy weakness in resource limited settings. Practical validation Independent embedded sensor Hardware validation is required on the embedded sensor platforms to further ensure that the computation is feasible, and energy modeling correct under realistic deployment requirements. Also, the framework should be expanded to include multi-hop wireless sensor network topology with dynamic routing and cooperative aggregation as a valuable direction in enhancing the scalability and resilience of large scale implementations. The general impact of this work is that it provides a principled and scalable contribution to sustainable edge intelligence of energy constrained wireless sensors networks.

References

1. Ahmed, E., & Rehmani, M. H. (2017). Mobile edge computing: Opportunities, solutions, and challenges. Future Generation Computer Systems, 70, 59–63.
2. Bonawitz, K., Eichner, H., Grieskamp, W., Huba, D., Ingerman, A., Ivanov, V., ... Roselander, J. (2019). Towards federated learning at scale: System design. Proceedings of Machine Learning and Systems, 1, 374–388.
3. Bonawitz, K., Ivanov, V., Kreuter, B., Marcedone, A., McMahan, H. B., Patel, S., ... Seth, K. (2017). Practical secure aggregation for privacy-preserving machine learning. In Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security (pp. 1175–1191). ACM.
4. Chiang, M., & Zhang, T. (2016). Fog and IoT: An overview of research opportunities. IEEE Internet of Things Journal, 3(6), 854–864.
5. Kairouz, P., & McMahan, H. B. (2021). Advances and open problems in federated learning. Foundations and Trends in Machine Learning, 14(1–2), 1–210.
6. Karimireddy, S. P., Kale, S., Mohri, M., Reddi, S., Stich, S. U., & Suresh, A. T. (2020). SCAFFOLD: Stochastic controlled averaging for federated learning. In Proceedings of the International Conference on Machine Learning (pp. 5132–5143). PMLR.
7. Li, T., Sahu, A. K., Zaheer, M., Sanjabi, M., Talwalkar, A., & Smith, V. (2020). Federated optimization in heterogeneous networks. Proceedings of Machine Learning and Systems, 2, 429–450.
8. Mao, Y., You, C., Zhang, J., Huang, K., & Letaief, K. B. (2017). A survey on mobile edge computing: The communication perspective. IEEE Communications Surveys & Tutorials, 19(4), 2322–2358.
9. McMahan, H. B., Moore, E., Ramage, D., Hampson, S., & Agüera y Arcas, B. (2017). Communication-efficient learning of deep networks from decentralized data. In Proceedings of the International Conference on Artificial Intelligence and Statistics (pp. 1273–1282). PMLR.
10. Nishio, T., & Yonetani, R. (2019). Client selection for federated learning with heterogeneous resources in mobile edge. In Proceedings of the IEEE International Conference on Communications (ICC) (pp. 1–7). IEEE.
11. Wang, J., Liu, Q., Liang, H., Joshi, G., & Poor, H. V. (2020). Tackling the objective inconsistency problem in heterogeneous federated optimization. In Advances in Neural Information Processing Systems (Vol. 33, pp. 7611–7623).
12. Wang, S., Tuor, T., Salonidis, T., Leung, K. K., Makaya, C., He, T., & Chan, K. (2019). Adaptive federated learning in resource constrained edge computing systems. IEEE Journal on Selected Areas in Communications, 37(6), 1205–1221.