IMPROVING PERFORMANCE OF SMALL-WORLD FEDERATED LEARNING SYSTEMS THROUGH CURATED MECHANISM OF NODE CONNECTION

Saket
Prof. Rajesh Hegde

<u>CONTENTS:-</u>

AIM:-

To design and implement a decentralized federated learning (DFL) system based on a small-world network topology that achieves enhanced accuracy and faster convergence compared to existing Watts-Strogatz and Kleinberg models. This is accomplished by introducing a customized connection mechanism that dynamically determines the probability of connection between nodes based on curated features.

Abstract:-

Federated Learning (FL) enables collaborative model training across multiple decentralized devices without sharing raw data, preserving privacy while distributing computational workloads. Decentralized Federated Learning (DFL) further removes the central server, relying on peer-to-peer communication within an overlay network. In such systems, the topology of the communication network significantly impacts model performance, especially in terms of accuracy and convergence speed.

This project focuses on improving the performance of DFL systems by modifying the underlying small-world network architecture. Traditional models such as Watts-Strogatz and Kleinberg provide useful small-world properties like high clustering and short path lengths but are limited by static connection rules. To address this, we propose a curated node connection mechanism that dynamically determines the probability of edge formation between nodes based on features such as model similarity, accuracy gain, node centrality, and data distribution.

Our proposed system demonstrates superior convergence and improved accuracy compared to standard small-world topologies. Extensive simulations on benchmark datasets validate the effectiveness of the curated connection strategy in enhancing the efficiency and robustness of decentralized learning networks.

INTRODUCTION:-

Federated Learning (FL) has emerged as a transformative approach to training machine learning models in a privacy-preserving and distributed manner. In FL, multiple clients collaboratively train a shared model while keeping their local data private. This technique is particularly useful in scenarios where data sharing is restricted due to privacy concerns or regulatory constraints, such as in healthcare, finance, and mobile devices.

Traditionally, FL relies on a centralized server to aggregate models from all participating clients. However, such centralized architectures suffer from single points of failure, potential scalability limitations, and communication bottlenecks. To overcome these issues, Decentralized Federated Learning (DFL) has gained traction. In DFL, the aggregation and communication processes are distributed across a network of peer nodes, eliminating the need for a central coordinator. This makes the system more robust, scalable, and fault-tolerant.

The topology of the communication network in DFL plays a crucial role in determining the efficiency and effectiveness of model training. Small-world network models—such as the Watts-Strogatz and Kleinberg graphs—have been widely adopted due to their ability to maintain short communication paths while preserving high local clustering. These properties help improve information dissemination and convergence speed in decentralized environments.

However, the connection strategies in standard small-world models are often static and structure-based, ignoring critical node-level dynamics like model quality, data similarity, or performance history. This can lead to inefficient information flow, slow convergence, or suboptimal accuracy in non-IID or heterogeneous environments.

To address these limitations, this project introduces a curated node connection mechanism that adapts dynamically to evolving network and model conditions. By calculating connection probabilities based on multiple node-level features—such as validation accuracy, accuracy gain, centrality, model similarity, and data distribution—we aim to build a feature-aware small-world topology that fosters faster and more accurate decentralized learning.

Through empirical simulations and comparative analysis with standard Watts-Strogatz and Kleinberg-based DFL systems, we demonstrate that our proposed approach yields significant improvements in convergence speed and final model accuracy, while maintaining communication efficiency.

Methodology:-

This project proposes a novel decentralized federated learning (DFL) framework built on a custom small-world topology where node connections are formed dynamically based on multiple learning and network factors. Unlike traditional Watts-Strogatz or Kleinberg models, which rely on fixed probabilities or distance-based shortcuts, our approach computes the edge formation probability using a feature-aware function.

Each node in the network is assumed to train a local model on non-IID data. At the end of each communication round, nodes evaluate their performance and selectively exchange model parameters with other nodes based on the computed connection probability. This probability is not fixed but varies dynamically based on local accuracy, accuracy gain, model similarity, and grid distance.

The key innovation lies in the Edge Probability Function, which combines:

Model performance indicators (accuracy, accuracy gain)

Structural graph proximity (Manhattan distance)

Model similarity (cosine similarity between weight vectors)

Tunable hyperparameters to control the influence of each factor

This allows the network topology to evolve adaptively, promoting communication between high-performing and similar nodes, while preserving the small-world property for efficient information propagation.

Mathematical Analysis:-

Let the probability of forming an edge between nodes i and j at round r be denoted as $P_{ij}^{(r)}$. We define probability as:

$$P_{ij}^{(r)} = \alpha \cdot e^{-\lambda|a_i^{(r)} - a_j^{(r)}|} \cdot (\Delta a_i^{(r)} + \Delta a_j^{(r)}) \cdot (\beta \cdot \mathrm{sim}_{ij}) \cdot \frac{1}{d_{ij}^r}$$

Where:

- $a_i^{(r)}$: Local accuracy of node i at round r
- $\Delta a_i^{(r)} = a_i^{(r)} - a_i^{(r-1)}$: Change in accuracy (i.e., accuracy gain)
- $\mathbf{w}_i^{(r)}$: Weight vector of node i at round r
- $\mathrm{sim}_{ij} = \cos(\mathbf{w}_i^{(r)}, \mathbf{w}_j^{(r)})$: Cosine similarity between model weights
- d_{ij}: Manhattan distance between nodes i and j in the network grid
- r: Kleinberg exponent (controls long-range connection decay)
- α, λ, β: Tunable positive constants

INTERPRETATION OF TERMS:-

- **Similarity & Performance Matching:**
  The exponential term $e^{-\lambda|a_i^{(r)} - a_j^{(r)}|}$ favors connections between nodes with similar validation accuracy, encouraging consensus among peers of comparable performance.
- **Learning Progress Incentive:**
  The term $\Delta a_i^{(r)} + \Delta a_j^{(r)}$ rewards nodes that are actively learning or improving, ensuring that highly dynamic and progressing nodes are more connected.
- **Model Similarity Weighting:**
  The cosine similarity term sim_{ij} ensures structural alignment between models being exchanged, improving aggregation quality.
- **Distance-Based Penalty:**
  The inverse polynomial term $\frac{1}{d_{ij}^r}$ preserves the small-world characteristic by penalizing long-distance connections, similar to the Kleinberg model.

Tunable Parameters:
α: Scales the overall connection probability.

λ: Controls how sharply connection probability decays with accuracy difference.
β: Balances the weight of cosine similarity.
r: Controls the spatial locality (higher r makes long-range links rarer).

This formulation enables the construction of a feature-curated, performance-sensitive, small-world DFL network that adapts over time to improve convergence and accuracy while respecting the locality and efficiency constraints of decentralized systems.

IMPLEMENTATION DETAILS:-

The proposed system was implemented as a simulation of decentralized federated learning (DFL) using a feature-aware small-world topology. The goal was to evaluate how curated, dynamic connection probabilities between nodes can improve model convergence and accuracy over traditional topologies like Watts-Strogatz and Kleinberg.

1. Tools and Frameworks Used

Programming Language: Python

Key Libraries:

- NumPy and SciPy: Numerical computation, linear algebra
- NetworkX: Graph generation and manipulation
- scikit-learn: Dataset loading and local model training (SGDClassifier)
- Matplotlib: Result visualization

2. Dataset

PENDIGITS dataset from OpenML and MNIST was used as a benchmark for classification tasks in a non-IID setting.

Each node received a non-identically distributed partition of the dataset to reflect real-world data heterogeneity.

3. Node Simulation and Local Training
Each node independently trains a local model using stochastic gradient descent (SGD) on its assigned data.

At the end of each communication round, nodes evaluate their local accuracy on a shared validation set.

Accuracy change is tracked across rounds for dynamic edge computation.

4. Network Topology Initialization
Nodes are placed on a 2D grid for calculating Manhattan distance.
An initial local ring or grid-based connectivity is established.
The curated small-world edges are formed dynamically at each round using the proposed edge probability function.

5. Dynamic Edge Formation (Per Round)
For each pair of nodes compute:

Local accuracy, accuracy gain and cosine similarity between weight vectors.
Use the edge probability function to sample edges probabilistically.
Construct the dynamic adjacency list based on sampled edges.

6. Communication and Model Aggregation

Each node communicates only with its current neighbors (based on the dynamic graph).

Aggregation is done via weighted averaging of model parameters using neighbor similarities or uniform weights.

Optionally, multiple consensus steps are allowed per round for improved information mixing.

7. Evaluation Metrics

Validation Accuracy: Measured at each round to track global learning progress.

Convergence Speed: Measured by rounds needed to reach a certain accuracy threshold.

Topology Metrics: Average clustering coefficient and path length of the dynamic graph.

8. Baseline Comparison

The curated model was compared against:

Watts-Strogatz (WS) topology with fixed rewiring probability.

Kleinberg small-world topology with static long-range links based on distance decay.

No communication and fully-connected extremes for benchmarking lower and upper performance bounds.

CODE SNIPPET:-

```
def kleinberg_neighbors_dynamic(n_nodes, distance_matrix, centrality,
sim_data, grads, accs, acc_gains,
                                 n_links=2, random_state=None):
    rng = np.random.default_rng(random_state)
    probs_matrix = np.zeros((n_nodes, n_nodes))

    for i in range(n_nodes):
        for j in range(n_nodes):
            if i != j:
                d = distance_matrix[i, j]
                sim = sim_data[i, j]
                grad = grads[i, j]
                acc = accs[j]
                gain = acc_gains[j]
                cen = centrality[j]

                prob = (1 / d) * sim * grad * acc * gain * cen
                probs_matrix[i, j] = prob

    for i in range(n_nodes):
        row = probs_matrix[i]
        row_sum = row.sum()
        if row_sum > 0:
            probs_matrix[i] = row / row_sum
        else:
            row[:] = 0
            row[np.arange(n_nodes) != i] = 1 / (n_nodes - 1)

    neighbors = []
    for i in range(n_nodes):
        row = probs_matrix[i]
        candidates = np.where(row > 0)[0]
        p = row[candidates]
        p /= p.sum()
        selected = rng.choice(candidates, size=min(n_links, len(candidates)),
replace=False, p=p)
        neighbors.append(list(selected))

    return neighbors
```

<u>RESULTS:-</u>

The above results are measured against the Pendigits Dataset with the custom model converging faster and with significant more accuracy than the existing models.

SN NO	N=20 ACCURACY	N=20 CONVERGENCE SPEED	N=50 ACCURACY	N=50 CONVERGENCE SPEED	N=100 ACCURACY	N=100 CONVERGENCE SPEED
WATTS-STROGATZ MODEL	90.21	28	89.91	52	87.1	28
NEWMAN MODEL	91.64	65	87.68	41	87.04	27
SCALE-FREE MODEL	93.8	100+	93.1	100+	92.1	100+
CUSTOM MODEL	98.39	12	98.69	13	97.61	8

Similarly, measured on a MNIST dataset, we have similar inferences:-

SN NO	N=20 ACCURACY	N=20 CONVERGENCE SPEED	N=50 ACCURACY	N=50 CONVERGENCE SPEED	N=100 ACCURACY	N=100 CONVERGENCE SPEED
WATTS-STROGATZ MODEL	91.08	28	94.08	51	96.71	31
NEWMAN MODEL	92.04	21	95.11	25	95.19	15
SCALE-FREE MODEL	96.01	33	96.03	38	94.87	20
CUSTOM MODEL	98.29	23	98.11	29	98.01	11

INFERENCES:-

The experimental evaluation of the proposed custom small-world decentralized federated learning (DFL) model on the Pendigits dataset demonstrates clear improvements over existing network models, including Watts-Strogatz, Newman, and Scale-Free architectures. The results were evaluated for three different network sizes: N = 20, N = 50, and N = 100. The key inferences are summarized below:

1. Superior Accuracy Across All Network Sizes
The custom model consistently achieves the highest validation accuracy across all network sizes:

N = 20: 98.39%

N = 50: 98.69%

N = 100: 97.61%

Compared to Watts-Strogatz and Newman models, the accuracy improvement is substantial (up to +8% in some cases).

Even against the Scale-Free model (which typically performs well due to hub-like connections), the custom model achieves ~5% higher accuracy at N=100, showing robust scalability.

2. Drastic Reduction in Convergence Time
The convergence speed (measured as the number of communication rounds needed to reach a stable accuracy) is significantly faster in the custom model:

N = 100: Only 8 rounds to converge, compared to 28 rounds in Watts-Strogatz and 100+ in the Scale-Free model.

This demonstrates that the adaptive edge probability function accelerates learning by intelligently connecting nodes based on learning dynamics, rather than static structural assumptions.

3. Weaknesses of Existing Models
The Watts-Strogatz model, while structured and local, lacks dynamic awareness of node performance, leading to slower convergence and slightly lower accuracy.

The Newman model, though more flexible in edge generation, still underperforms due to lack of data- and performance-based edge selection.

The Scale-Free model, despite producing high accuracy in some cases, suffers from excessive convergence time, making it unsuitable for resource-constrained or real-time DFL system.

4. Effectiveness of the Curated Connection Mechanism

The curated edge probability function that incorporates accuracy, accuracy gain, cosine similarity, and grid distance proves highly effective:

It strikes a balance between exploiting local clusters and exploring distant but similar nodes.

The resulting topology adapts over time to prioritize high-performing and improving nodes, promoting better information flow.

5. Scalability and Robustness

The performance gains of the custom model are consistent across small, medium, and large-scale networks.

This suggests that the proposed connection mechanism scales well and can generalize to a variety of DFL applications.

CONCLUSION:-

This project presents a novel approach to improving the performance of decentralized federated learning (DFL) systems through a curated, feature-aware node connection mechanism embedded within a small-world network architecture. By integrating learning dynamics—such as local accuracy, accuracy gain, model similarity, and grid distance—into the edge formation process, the proposed system intelligently adapts the communication topology over time.

Experimental results on the Pendigits dataset clearly demonstrate that the custom model outperforms traditional Watts-Strogatz, Newman, and Scale-Free network models in both convergence speed and accuracy across varying network sizes (N = 20, 50, 100). The model achieves convergence up to 10× faster than existing approaches and delivers up to 8% higher validation accuracy, proving the effectiveness of dynamic, performance-guided communication in DFL systems.

These findings validate the central hypothesis of the project: that customizing inter-node connectivity based on curated, multi-factor features can significantly enhance learning efficiency in decentralized systems. The proposed framework offers a promising step forward in making DFL more practical, adaptive, and performance-driven—particularly in non-IID, heterogeneous, or resource-constrained environments.