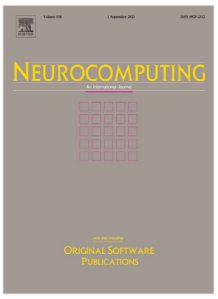
Journal Pre-proof



RFCSC: Communication efficient reinforcement federated learning with dynamic client selection and adaptive gradient compression

Zhenhui Pan, Yawen Li, Zeli Guan, Meiyu Liang, Ang Li, Jia Wang, Feifei Kou

PII: S0925-2312(24)01443-7
DOI: https://doi.org/10.1016/j.neucom.2024.128672
Reference: NEUCOM 128672

To appear in: *Neurocomputing*

Received date : 14 November 2023
Revised date : 31 August 2024
Accepted date : 29 September 2024

Please cite this article as: Z. Pan, Y. Li, Z. Guan et al., RFCSC: Communication efficient reinforcement federated learning with dynamic client selection and adaptive gradient compression, *Neurocomputing* (2024), doi: https://doi.org/10.1016/j.neucom.2024.128672.

RFCSC: Communication Efficient Reinforcement Federated Learning with Dynamic Client Selection and Adaptive Gradient Compression

Zhenhui Pan[a], Yawen Li[b,*], Zeli Guan[a], Meiyu Liang[a], Ang Li[a], Jia Wang[a] and Feifei Kou[a]

[a]Beijing Key Laboratory of Intelligent Telecommunication Software and Multimedia, School of Computer Science, Beijing University of Posts and Telecommunications, Beijing, 100876, China

[b]School of Economics and Management, Beijing University of Posts and Telecommunications, Beijing, 100876, China

ARTICLE INFO

Keywords:
Reinforcement federated learning
Dynamic client selection
Adaptive Gradient Compression

ABSTRACT

In the field of public safety, high-quality data is often owned by governments, companies, and organizations, making it difficult to train effective models through centralized datasets. This paper leverages federated learning as a mechanism to address data privacy concerns. Within the federated learning framework, the data across different clients is typically Non-IID (non-Independently and Identically Distributed). Furthermore, The complexity of sensitive image recognition tasks in public safety, along with the large number of model parameters, can lead to communication congestion issues in federated learning. To solve these challenges, this paper proposes a Reinforcement Federated Client Selection and Gradient Compression Method(RFCSC). By integrating client data prototypes with accuracy metrics, we dynamically assess the contribution of each client in federated learning. Through an intelligent dynamic incentive mechanism based on reinforcement learning, high-quality client nodes are dynamically selected to participate in federated learning, achieving dynamic adaptive aggregation, reducing the influence of Non-IID data, reducing communication cost, enhancing model accuracy, and achieving a balance between quality and efficiency. To address the issue of high communication cost for parameter transmission, we propose an adaptive model compression strategy in reinforcement federated learning which make local clients find an appropriate compression rate. This approach not only reduces gradient communication overhead but also minimizes the impact of compression on model accuracy. The effectiveness of our proposed approach are corroborated through comprehensive experiments conducted on one private dataset and two public datasets.

1. Introduction

In recent years, public safety emergencies have received widespread attention. Researchers are starting to use advanced deep learning techniques for detecting events that are sensitive to public safety. [1, 2]. This research area has become critically important, aiming to provide real-time safety alerts and protection for public areas. Leveraging deep learning-based detection methodologies, remarkable advancements have been realized in the sphere of public safety[3]. By employing deep neural network models, researchers can derive intricate and highly distinguishing features from visual data[4]. These advancements have improved detection systems' ability to understand visual cues and the context related to potential threats, thus significantly increasing their accuracy and reliability[5].

In the domain of public safety, a significant portion of relevant images is stored within corporate and institutional databases, leading to challenges in inter-departmental data sharing. Federated learning enables various departments to collaboratively train models through multi-party cooperation[6]. In federated learning, departments share models without sharing data, allowing for model training while protecting data privacy[7]. Another notable challenge in federated learning is communication efficiency[7]. As participants train locally and only send model updates to the central server, frequent data exchanges are required.

*Corresponding author
warmly0716@126.com (Y. Li)
ORCID(s):

For tasks with large datasets and complex models, this can lead to significant communication cost, leading to increased network latency and bandwidth pressure[8]. Therefore, integrating model compression techniques into federated learning becomes crucial to address this challenge. Model compression is a crucial technique aimed at reducing the model's size and the associated communication cost, addressing challenges related to communication efficiency and network latency[9]. Recognized as a vital technological advancement, model compression is designed to manage the demanding storage and computational needs inherent to deep learning architectures[10]. The primary goal is to decrease the model's parameter count and computational demands while maintaining its predictive accuracy[11]. This ensures faster deployment and inference of the model. The domain of model compression offers various methods, including network pruning, parameter quantization, low-rank matrix factorization, and knowledge distillation[12, 13, 14, 15]. However, these model compression methods may affect model accuracy[16]. Introducing reinforcement learning to select the optimal clients can enhance model precision.

The model compression methods may affect model accuracy[16]. Introducing reinforcement learning to select the optimal clients can enhance model precision. Reinforcement learning is a distinct branch of machine learning techniques, centered on optimizing decision-making strategies through the dynamic interaction between an agent and its environment[17]. This method has been widely applied in robot control, autonomous driving, and federated learning.[18, 19, 20]. There is a growing interest in using

reinforcement learning algorithms for selecting clients[20]. However, these methods use parameters reduced by PCA as the state space, resulting in an excessively large dimensionality of the state space, which limits the improvement of model accuracy.

The task of detecting public safety sensitive images is fraught with a constellation of challenges, each contributing to the complexity of the endeavor. Firstly, in real-world applications, public safety data typically belongs to governments, companies, or organizations. The proprietary nature of this data precludes its free dissemination among disparate parties, thereby complicating the training of specialized models for sensitive image detection. Secondly, in federated learning, a variety of devices from mobile to edge contribute updates from their local data. Since these datasets are Non-IID, combining these updates can introduce inconsistencies. These can reduce the overall effectiveness of the final global model. Thirdly, due to the complexity of public safety-sensitive image detection tasks, the large number of model parameters, and the substantial memory occupation, transmission time and latency are increased in federated learning.

To solve these challenges, this paper proposes a Reinforcement Federated Client Selection and Compression method based on Double Deep Q-Network (DDQN), referred to as RFCSC. To address the challenges of data diversity and uneven quality among participating nodes, RFCSC combines data prototypes with accuracy metrics to dynamically assess each client's contribution, intelligently selecting high-quality nodes to enhance model efficiency and performance. To tackle the issues of data heterogeneity and non-independent and Non-IID data among nodes, RFCSC implements an adaptive aggregation strategy within the federated learning framework, effectively reducing bias in global model aggregation, especially when handling complex tasks. To mitigate the communication overhead caused by frequent model updates, RFCSC constructs a reinforcement learning reward based on model accuracy and communication volume, dynamically adjusting client selection and compression strategies to reduce bandwidth requirements while maintaining model accuracy. DDQN is used to model the client selection strategy, updating the weights of the online network to optimize client selection based on rewards, thereby improving the Q-value for each client and enhancing the global model's performance and stability. The main contributions of this paper include the following three points:

1. We propose an efficient reinforcement federated learning framework that incorporates an intelligent federated dynamic incentive mechanism to dynamically select high-quality federated nodes for model adaptive aggregation. Furthermore, the framework utilizes a model adaptive compression strategy to reduce the communication cost in the federated network. By integrating gradient compression techniques, we reduce communication overhead, achieving a balance between performance quality and efficiency.
2. We propose a dynamic client selection strategy for federated learning based on deep reinforcement learning. By integrating client data prototypes with accuracy metrics, we dynamically assess the contribution of each client in federated learning. This has led to the creation of an intelligent incentive mechanism and a dynamic adaptive aggregation strategy within reinforcement federated learning, enhancing the balance between quality and efficiency in the federated learning process.
3. We propose an adaptive model gradient compression mechanism. The local model selects an appropriate compression rate based on the distinctiveness of the data, preserving valuable gradient information from the client models. This approach not only reduces gradient communication overhead but also minimizes the impact of compression on model accuracy.

The structure of this paper is as follows. Section 2 introduces the related work of this paper. In section 3, we introduces the proposed RFCSC algorithm. In section 4, we present the experimental results . We concludes this paper in section 5.

2. Related work

2.1. Public Safety Sensitive Event Detection

Public Safety Sensitive Event Detection has emerged as a pivotal domain within computer vision, with the primary objective of discerning sensitive content through meticulous image and video analyses[21, 22]. Studies in this realm predominantly leverage deep learning and computer vision methodologies[23, 24, 25, 26]. The typical workflow encompasses two cardinal phases: feature extraction and subsequent classification. Deep neural networks serve as the go-to architectures for extracting salient features, such as color, texture, and shape, from visual data. These extracted features subsequently undergo classification processes to ascertain the existence of sensitive content. Some researchers have begun leveraging advanced deep learning techniques to detect events related to public safety[27, 28]. For instance, Li et al. proposed a dynamic multi-objective optimization algorithm based on multi-strategy adaptive selection (MSAS-DMOA), which effectively addresses negative transfer issues in complex, dynamic environments, thereby enhancing the adaptability and performance of models[29]. While the advent of deep learning has undeniably ushered in transformative advancements in the domain, particularly in enhancing accuracy and robustness[30, 31, 32], Researchers have also adopted attention-based deep neural network models to extract complex and highly discriminative features from visual data. These models have shown significant success in fields such as medical image analysis and small object detection [33]. While existing methods have achieved commendable results in centralized datasets, there is still a lack of research related to the detection of public safety incidents under federated conditions.

2.2. Federated Learning

Federated learning is a significant research direction for addressing the problem of learning from distributed data. The aim of federated learning is to train local models on local data and merge them into a global model by exchanging parameters, thereby achieving centralized training without exposing the original data. The most representative method is FedAvg [34], which trains models on local devices and then sends the average of the model updates to a central server for aggregation to construct a global model. In this process, due to the variability of local training datasets, the model updates from each local device may exhibit certain biases and variances. Therefore, FedAvg employs a weighted averaging approach to ensure the accuracy and robustness of the global model. Unlike traditional centralized learning algorithms, FedAvg faces new challenges such as node selection, communication efficiency, data privacy, and security. To address these challenges, researchers have proposed various improved FedAvg algorithms, such as FedProx, FedAvgM, FedAdapt, and FedMA[35, 36, 37, 38]. These algorithms mainly improve the performance of FedAvg by adopting different strategies to solve issues like node selection, model compression and data privacy protection. However, these methods may require more communication rounds to achieve satisfactory performance, and they do not take into account the issue of Non-IID data.

2.3. Gradient Compression

Gradient compression emerges as a crucial technique tailored to reduce the magnitude of gradient transmission, addressing the communication overhead intrinsic to distributed machine learning frameworks, notably federated learning. A prominent strategy within the gradient compression arsenal is parameter quantization[39, 40]. This method optimizes the bit-width requisite for transmission by converting floating-point parameters into more compact, low-precision integer or fixed-point representations. Empirical studies have underscored the efficacy of these algorithms in significantly diminishing gradient transmission volume without compromising model accuracy. Another notable approach within gradient compression is gradient pruning[41, 42, 43], which leverages sparsification techniques to discard gradients that fall beneath a specified threshold, thereby curtailing the number of transmitted gradient values. There exists a spectrum of gradient pruning techniques, spanning from threshold-centric methods to sparse matrix-oriented strategies, all instrumental in curbing data transmission volume and, consequently, communication overhead. However, while the aforementioned methods can reduce communication overhead, they simultaneously impact the performance of the model.

2.4. Reinforcement Learning

Reinforcement Learning (RL) emerges as a foundational approach in training intelligent agents to autonomously make decisions, aiming to maximize cumulative rewards through interactions with their environment1. This technique has been adopted across a wide array of fields, encompassing robotic control, node selection in graph neural networks, and natural language processing[17, 18, 44]. Q-Learning, an RL algorithm rooted in value iteration, addresses decision-making challenges within the Markov Decision Processes (MDP) framework[45]. Central to this algorithm is the iterative optimization of a Q-value function, which subsequently informs action choices within a defined state space. DQN enhance the Q-Learning framework by integrating deep neural networks to approximate the Q-value function, thereby elevating the algorithm's precision and efficiency[46]. Within DQN, experience replay mechanisms store and stochastically sample previous interactions, facilitating the optimization of neural network parameters[47]. This randomized sampling strategy alleviates correlation challenges during training and optimizes sample efficiency. To further ensure stability, DQN incorporates a target Q-value network, mitigating fluctuations in target value estimations. The confluence of reinforcement learning and federated learning has also been explored, with RL being applied to cross-device and cross-domain federated transfer learning. This integration not only bolsters model adaptability and generalization but also reinforces security and privacy safeguards.

3. RFCSC: Overview of the Model Framework

In this study, we propose a Reinforcement Federated Client Selection and Compression (RFCSC) methodology based on DDQN. The approach utilizes federated learning for collaborative global model training without raw data sharing. We design local training and gradient compression modules for public safety event detection. To mitigate communication overhead, gradient compression techniques are employed to minimize data transmission. Traditional algorithms often assume independently and identically distributed (iid) data, an assumption frequently violated in federated settings. Non-iid conditions can lead to heterogeneous model updates when using FedAvg, affecting global model convergence. To address this, we incorporate reinforcement learning for optimal client selection each round. We define a reward function r based on accuracy improvement and consider factors like data distribution and quality. Data prototypes serve as the state variable s in the reinforcement learning algorithm, ensuring balanced client selection. The architecture is outlined in Fig. 1. The entire model framework comprises two parts: global aggregation and local training, corresponding to sections 3.2 and 3.3, respectively.

3.1. Local Model Learning of Public Safety Sensitive Event Detection

In federated learning, we employ a deep learning model for detecting safety-sensitive images. The model takes images or video frames as input and outputs safety threat event detection. It learns visual features to distinguish between

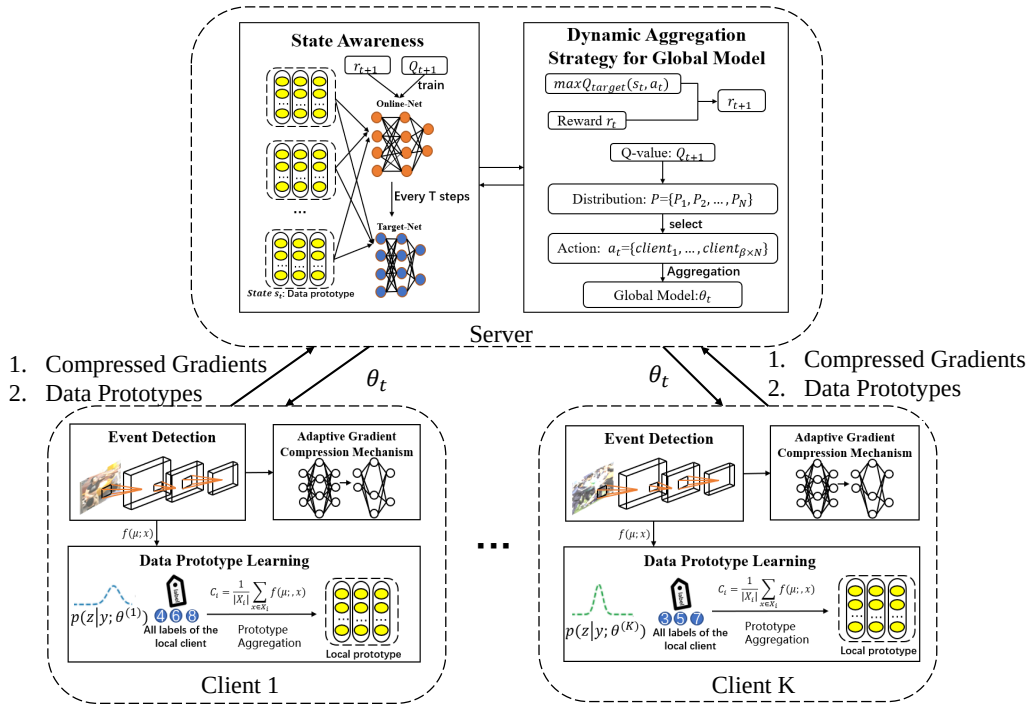


Figure 1: Overview of our Reinforced Federated Client Selection and Compression Method Based on DDQN.

safe and hazardous behaviors. The training involves preprocessing, forward and backward propagation, and parameter optimization. A chosen loss function measures the output-label discrepancy, and backpropagation updates the model parameters. A classifier attached to the output layer uses the Softmax function for final classification, as in Eq.(1)

$$p_{i,m} = \text{softmax}(W_{\text{client}} \cdot x_{i,m} + b_{\text{client}}) \tag{1}$$

Here, $p_{i,m}$ is the event classification probability, W_{client} and b_{client} are the trainable parameters of the local model, while $x_{i,m}$ represents the input images or video frames. The loss function employs cross-entropy to quantify the discrepancy between the model's output and the true labels, as indicated in Eq.(2):

$$L_1 = -\frac{1}{N}\sum_{i=1}^{N}\sum_{m=1}^{M} y_{i,m}\log(p_{i,m}) \tag{2}$$

where L_1 is the loss function of local model. N is the number of training samples, M is the number of sensitive behavior categories, and $y_{i,m}$ is the label of the i^{th} sample.

3.2. Federated Client Dynamic Adaptive Selection Strategy based on Deep Reinforcement Learning

This paper constructing an intelligent dynamic incentive mechanism for reinforcement federated learning. In order to better select the optimal clients that participate in the global model update, this paper employs deep reinforcement learning based on DDQN, using data prototypes and model accuracy as evaluation indicators for the selection of high-quality clients. The Double Deep Q-Network (DDQN) is used in federated learning to dynamically select clients for participation in training. It is implemented with two deep neural networks: the online network, which is responsible for selecting actions, and the target network, which calculates the target Q-values. The Q-network architecture consists of two fully connected layers, with the input being the feature representation of the state and the output being the Q-values for each possible action. ReLU activation functions are employed to enhance the non-linear capacity of the network. An experience replay buffer is used to store past experiences, and its size should be sufficiently large to cover a variety of scenarios. The update frequency determines how often the online network is sampled and updated from the buffer. The parameters of the target network are updated only after a certain number of steps to maintain training stability. Within the federated learning framework, DDQN predicts the Q-values for each client to select high-quality clients for model updates, thereby optimizing the efficiency and accuracy of global model training. This approach is particularly effective in scenarios with high client data heterogeneity. The design of states, actions, and rewards is as follows.

State in Reinforcement Federated Learning: Traditional Reinforcement Learning client selection employs network

parameters as the state, which can result in an excessively large state space. In RFCSC, we choose to use data prototypes[48] as the state space $s = \{C^1, C^2, \cdots, C^N\}$, where M is the number of labels, and C^i is the average prototype for i^{th} label. Our approach simultaneously considers data prototypes and accuracy as indicators. A data prototype is an intermediate representation in the model, representing a feature vector obtained after input data is processed through convolutional and pooling layers, followed by a fully connected layer. This feature vector can be viewed as a compact representation of the input data, encapsulating its significant features and patterns. The calculation of the data prototype ensures that it captures the essential characteristics of the input data, allowing it to effectively represent these features in subsequent model training and inference processes. In reinforcement learning or federated learning, the data prototype can serve as a feature representation of client data, which is useful for optimizing client selection strategies or the model updating process. The clients' data prototypes are obtained as shown in Eq.(3):

$$\begin{cases} C_i = \frac{1}{|X_i|} \sum_{x \in X_i} f(\mu; x) \\ C^k = \{C_1, C_2, \cdots, C_M\} \end{cases} \tag{3}$$

where x is the data of label i, $|X_i|$ is the number of label i, and μ is the parameter of the function f. M is the number of labels in client k. C^k is the data prototypes of client k. The state on server is: $s_t = \{C^1, C^2, \cdots, C^N\}$. N is the number of clients.

Action in Reinforcement Federated Learning: Given the current state, the Q-values for all clients are computed using the online network:

$$Q_i = Q_{online}(C^i, a_t; \theta_{online}) \tag{4}$$

where Q_{online} is online network. Q_i is the Q-value of client i. θ_{online} is the parameter of online network. The Q-values for all clients are subjected to a softmax operation to obtain a probability distribution over the clients:

$$P_i = \frac{e^{Q_i}}{\sum_j e^{Q_j}} \tag{5}$$

By utilizing these probabilities, clients are selected such that those with higher Q-values have a greater likelihood of being chosen, while clients with lower Q-values still retain a probability of selection. This approach ensures a balanced consideration of all clients regardless of their individual Q-values. The action is: $a_t = \{Client_1, Client_2, \cdots, Client_{\beta \times N}\}$. β is selected rate. N is the number of clients.

Reward in Reinforcement Federated Learning: In the context of reinforcement learning, we employ a reward mechanism that is intricately tied to the incremental changes in model accuracy. Specifically, the agent is endowed with a positive reward commensurate with the magnitude of improvement in model accuracy. This design paradigm serves to incentivize the agent to continually strive for enhancements in predictive accuracy in subsequent decision-making epochs. Such positive reinforcement facilitates the learning of more efficacious policies and behaviors, thereby contributing to the overall performance optimization of the global model. Conversely, if the model's accuracy experiences a decrement, the agent is subjected to a negative reward, equivalent to the negative value of the decline in accuracy. This architectural choice provides the agent with adverse feedback, signaling that its recent actions have led to a deterioration in model performance. Negative rewards act as a deterrent, discouraging the agent from engaging in actions that are detrimental to accuracy, and encouraging the pursuit of more optimal strategies for performance improvement, as illustrated in Eq.(6):

$$R(s,a) = \begin{cases} acc - acc_{\text{pre}}, & \text{if } acc > acc_{\text{pre}} \\ acc_{\text{pre}} - acc, & \text{if } acc < acc_{\text{pre}} \end{cases} \tag{6}$$

$R(s,a)$ is reward. Acc and acc_{pre} are this round's accuracy and last round's accuracy respectively. Let r_{t+1} represents the target reward, the calculation of which is detailed in Eq.(7):

$$r_{t+1} = r_t + \gamma \max_{a'} Q_{target}(s'_i, a'_i; \theta_{target}) \tag{7}$$

where r_t represents the reward at time t, γ is the discount factor, s'_i is the subsequent state, and θ_{target} are the parameters of the target network. Through iterative training, DDQN can progressively optimize the Q-value function and select the optimal action based on this Q-value function. This enables the agent to learn the best policy for taking actions in different states, thereby enhancing the model's accuracy.

In Experience Replay, we maintain an experience buffer to store the agent's experiences (states, actions, rewards, next states). A batch of samples is then randomly drawn from this buffer for training, breaking the correlation between samples. The fixed Target Network is used to compute the target Q-values, reducing the volatility of the targets. The parameters of the DDQN are updated using gradient descent to minimize the mean squared error between the predicted Q-values and the target Q-values. The loss function is described in Eq.(8):

$$L(\theta_{online}) = \frac{1}{N} \sum_i (r_{t+1} - Q(s_i, a_i; \theta_{online}))^2 \tag{8}$$

$$Q(S_t, a_t) = Q(S_t, a_t) + \alpha \Big[R_{t+1} \tag{9}$$

$$+ \gamma \max_a Q(S_{t+1}, a) \tag{10}$$

$$- Q(S_t, a_t) \Big] \tag{11}$$

where $L(\theta_{online})$) is the loss function, N is the batch size of the samples. We employ the backpropagation algorithm

Algorithm 1: Local Model Learning of Public Safety Sensitive Event Detection

Input: Client Data $\boldsymbol{X}_i$, Global model parameters θ, Compression ratio r , Epochs E, accuracy Acc_i

Output: Compressed gradient g_j prototype $\left\{C_1^i, C_2^i, \cdots, C_M^i\right\}$

for $t = 1, 2, \cdots, E$ **do**
  Sample a mini-batch B from $\boldsymbol{X}_i$
  compute the predictions $\hat{y} = f(\theta; B)$
  compute loss L using Eq.(2)
  compute the gradient $\nabla L(\theta; B) = \frac{\partial L}{\partial \theta}$
  Update the parameters: $\theta \leftarrow \theta - \eta \nabla L(\theta; B)$
end
Compute the Compressed gradient g_j by Eq.(10)
Compute the Prototype $\left\{C_1^i, C_2^i, \cdots, C_M^i\right\}$ by Eq.(3)
Compute the test accuracy Acc_i

to update the weights θ_{online} of the online Q-network. The target Q-network updates the weights every K steps:

$$\theta_{target} \leftarrow \theta_{online} \tag{12}$$

The primary advantage of DDQN is that by decoupling the action selection from the Q-value computation, it mitigates the overestimation of Q-values, leading to a more stable and accurate estimation of Q-values.

3.3. Adaptive Model Gradient Compression Mechanism

During the model training process, a high compression rate means that the amount of gradient data transmitted to the server is reduced. While this can significantly lower communication costs, it may also result in the loss of gradient information that is critical for model updates. These subtle yet important pieces of information may contain valuable hints about the direction of model convergence. If the compression rate is too high, such information might be overlooked, leading to reduced learning effectiveness and decreased accuracy of the model. To address this problem, the adaptive compression strategy aims to minimize the negative impact on model accuracy by dynamically adjusting the compression rate. In each training iteration, the strategy adjusts the compression rate based on the changes in the current model's accuracy. If the model's accuracy significantly declines under a high compression rate, the adaptive strategy reduces the compression rate to ensure more crucial gradient information is transmitted, thereby maintaining the model's learning effectiveness and accuracy. Conversely, if the model's accuracy remains stable or improves at a certain compression rate, the adaptive strategy attempts to increase the compression rate to further reduce communication cost. To reduce communication cost, we employ the Top-k sparsification technique to compress the

Algorithm 2: Server Dynamic Adaptive Aggregation of the Global Model.

Input: Set of clients N
Compressed gradient $G_N = \left\{g_1, g_2, \cdots, g_N\right\}$
Client's test Accuracy $\left\{Acc_1, Acc_2, \cdots, Acc_N\right\}$
Client's Prototype $\left\{C^1, C^2, \cdots, C^N\right\}$
DDQN networks $Q_{\text{online}}, Q_{\text{target}}$
Communication epoch E_s

Output: global model θ_t

for $t = 1, 2, \cdots, E_s$ **do**
  Observe the state s_t
  Select an action a_t(i.e., a set of clients for model update) based on the policy derived from $Q_{\text{online}}(s_t, \cdot)$
  The server updates global model using Eq.(13)
  Observe the new state $s_{t+1} = \left\{C^1, C^2, \cdots, C^N\right\}$
  Observe the new accuracy and set $acc = acc_{\text{new}}$
  Compute the reward r_t by using Eq.(6)
  Store transition (s_t, a_t, r_t, s_{t+1}) in replay buffer
  Update Q_{online} using Eq.(8)
  Every T steps, reset $Q_{\text{target}} = Q_{\text{online}}$
end
return global model θ_t

gradients of the model. Given a gradient vector $g \in \mathbb{R}^n$ of a model, we first compute the absolute value of each of its elements and sort them. Let $|g|_{(1)} \geq |g|_{(2)} \geq \cdots \geq |g|_{(n)}$ represent the descending order of the absolute values of the gradients. By real-time monitoring of model performance metrics, such as accuracy and loss values, the compression rate is dynamically adjusted.

$$g_i = \begin{cases} g_i & \text{if } |g_i| \geq |g|_{(k)} \\ 0 & \text{otherwise} \end{cases} \tag{13}$$

Upon applying the compression process, the resultant gradient vector retains merely k non-zero components, thereby yielding a compression efficacy of $\frac{n-k}{n} \times 100\%$. To achieve model quality and efficiency optimization, we employ an adaptive compression rate strategy. This strategy initiates with a predetermined compression rate and proceeds with model training and evaluation. If the model's performance exceeds a set threshold, the compression rate is incrementally increased; conversely, if the performance falls below the threshold, the compression rate is reduced accordingly. This process continues until a specified number of training iterations or other termination criteria are met. Ultimately, the model obtained using the optimal compression rate is deployed, ensuring effective model compression while maintaining high performance. The adaptive compression strategy using the following formulas:

$$CR_{\text{new}} = \begin{cases} CR_{\text{current}} + \Delta CR & \text{if } P_{\text{current}} > P_{\text{threshold}} \\ CR_{\text{current}} - \Delta CR & \text{if } P_{\text{current}} < P_{\text{threshold}} \\ CR_{\text{current}} & \text{otherwise} \end{cases} \quad (14)$$

where CR_{new} is the updated compression rate, CR_{current} is the current compression rate, ΔCR is the incremental adjustment value, and $P_{\text{threshold}}$ is the set performance threshold. By employing adaptive compression mechanism, we aim to identify the optimal compression rate, thereby sparsifying the uploaded gradients. This approach is intended to achieve a balance between quality and efficiency, optimizing the trade-off to ensure optimal performance.

3.4. Dynamic Aggregation Strategy for Global Model.

Q_i is the Q-value of client i computed by Eq.(4) which is the contribution of each client learned through reinforcement learning. Based on the contribution of clients and model accuracy, achieve dynamic adaptive aggregation. Only the selected clients upload gradients for aggregation, and the clients chosen vary each time. The update of the global gradient, denoted as $\overline{\nabla}_t$, is formulated as in Eq.(12):

$$\overline{\nabla} t = \frac{1}{N_t} \sum_{n=1}^{N_t} g_j \quad (15)$$

where N is the number of clients, β is select rate, $N_t = \beta \times N$, N_t is the number of clients selected dynamically at time t based on the dynamic aggregation strategy. The global model is updated as in Eq.(13):

$$\theta_t = \theta_{t-1} - \eta_1 * \overline{\nabla}_t \quad (16)$$

where θ_t represents the global model parameters after the t^{th} aggregation, and η_1 is the learning rate. This dynamic and adaptive client selection strategy enhances the flexibility and efficiency of federated learning, enabling it to better cope with various uncertainties and changes. The local training process is illustrated in Algorithm 1. The dynamic aggregation process, which determines N_t at each time step, is illustrated in Algorithm 2.

4. Experimental Results and Analysis

4.1. Dataset

In our experiments, we employed three datasets for testing and validation. The selection of these three datasets is attributed to their diversity, which aids in verifying the generalization performance of our proposed method across varied tasks and data distributions.

Violence Image Recognition (VIR) [49]: The dataset pertains to behaviors endangering public safety. It encompasses a variety of image types, such as physical altercations and assaults. The original dataset is a binary classification image dataset. By incorporating additional data gathered through our research, we expanded the dataset from 2 to 4 categories. This enhancement ensures a more comprehensive representation, capturing a broader spectrum of behaviors jeopardizing public safety as observed in the real world.

CIFAR-10[50]:The CIFAR-10 is a prevalent computer vision dataset, comprising 60,000 color images of 32x32 resolution, distributed across 10 categories with each category containing 6,000 images. This dataset is extensively employed for image recognition tasks and stands as a benchmark dataset highly esteemed in both the academic and industrial communities.

FEMNIST[51]: The FEMNIST dataset is an extension of the MNIST dataset, encompassing approximately 700,000 images of handwritten characters. Distinct from MNIST, FEMNIST not only includes digits but also encompasses both uppercase and lowercase letters, rendering the tasks associated with this dataset more challenging.

In terms of data partitioning, the Dirichlet partitioning[52] is a prevalent strategy employed in federated learning scenarios. This strategy, grounded in the Dirichlet distribution, is adept at generating diverse non-uniform data distributions. The shape and scale of this distribution are governed by what is termed the concentration parameter. A principal challenge associated with Dirichlet partitioning is the selection of an appropriate concentration parameter to engender a data distribution reflective of the actual problem at hand. Improper parameter choices might culminate in a data distribution that substantially deviates from the real-world problem, thereby impinging upon the performance of the federated learning algorithm. Consequently, we adopt a label-based data partitioning strategy. For instance, a parameter $\alpha = 0.8$ signifies that 80% of the data within a client pertains to a single label, while the residual 20% corresponds to other labels. Such label-centric partitioning adeptly simulates skewed data distributions, enabling the model to account for this imbalance during training, thereby enhancing its generalization capabilities and robustness. With label-based partitioning, one can explicitly control the degree of data skewness by adjusting the label proportions, whereas the distributions generated by Dirichlet partitioning are less intuitively interpretable.

4.2. Evaluation Metrics and Experimental Configuration

This paper uses test accuracy, the number of communication rounds, and the amount of parameters uploaded by clients as the main evaluation metrics. Accuracy measures the model's classification performance on different datasets (FEMNIST, CIFAR-10, and VIR), reflecting the effectiveness of each method under various non-independent and Non-IID conditions and different compression rates. The number of communication rounds evaluates the training rounds required for each method to reach a specific accuracy target, thereby indicating the convergence speed of the method. The amount of parameters uploaded by

clients reflects the communication cost for each method to achieve the same accuracy goal. Through comparative analysis of these metrics, we can comprehensively evaluate the performance of the proposed method (RFCSC) under different experimental conditions, particularly its advantages in improving model accuracy and reducing communication costs. In our experimental framework, we adopted the Stochastic Gradient Descent (SGD) optimization technique, configuring the momentum and weight decay parameters to 0.9 and 0.0001, respectively. The learning rate η_1 was systematically varied within the interval $[0.001, 0.05]$. For the reinforcement learning paradigm governing client selection, we employed the Adam optimizer, characterized by a momentum of 0.9 and a learning rate η_2 of 0.0001. Within the DDQN methodology, we designated a discount factor of 0.95 and initialized the exploration rate to unity. To ensure comprehensive model convergence, epochs were set to 300 for both the FEMNIST and CIFAR-10 datasets, while the VIR dataset being restricted to 100.We use α to denote as the imbalanced Non-IID of dataset. The values are set to 0.7, 0.8 and 0.9. Client selection rates were fixed at 0.2 and 0.4. If client number is 50 and client selection rate is 0.2 means that server will choose 10 clients. It is noteworthy that a consistent hyperparameter configuration was maintained across all client-side computations during the federated learning phase.

We found the optimal compression rates for each dataset through an adaptive compression mechanism. For the VIR dataset, we leveraged parameters from a pre-trained Resnet50 architecture, the best compression rate is 0.5. Moreover, for the CIFAR-10 and FEMNIST datasets, we utilized a Convolutional Neural Network (CNN) initialized with random weights. The best of compression rate for these two datasets is 0.4.

The computational framework was instantiated using the PyTorch library, executed on a GPU configuration comprising 2 NVIDIA RTX2080TI units. Furthermore, an exhaustive search within the set $\{0.01, 0.0075, 0.005, 0.001\}$ facilitated the identification of the optimal learning rate for benchmark comparisons.

4.3. Baselines

To thoroughly test the effectiveness of our reinforcement-based federated learning method, we compared it with several key federated learning algorithms in our experiments. It is imperative to note that throughout these comparative evaluations, we maintained uniformity in terms of the dataset, model architecture, and experimental configurations.

E3CS[53]: The E3CS algorithm meticulously addresses the issue of data heterogeneity across clients. E3CS employs a probabilistic allocation strategy, designating a foundational probability to each client and subsequently distributing the residual probability contingent on weights.

Pow-d[54]: The Pow-d client selection algorithm diverges from conventional methods such as random selection or those based on data skewness. This algorithm exhibits a predilection for clients with elevated local losses, thereby facilitating accelerated error convergence.

Cluster-sample[55]: The Cluster-sample methodology encompasses two distinct algorithms. In Algorithm 1, for every client, it delineates a set quantity that must be represented with a specified probability, ensuring uniform representation of each client across all distributions. Algorithm 2 employs clustering to initialize allocations, the largest sets are then considered for the allocation of any remaining samples.

DivFL[56]:In each global iteration, DivFL identifies a subset whose aggregated model updates approximate the collective updates from all clients. This approach aims to constrain the variance introduced by subset selection in model updates, thereby accelerating the convergence rate of the model's learning process.

FedAvg[34]: In each global iteration, FedAvg aggregates model updates from a subset of clients to compute a global model update. This approach aims to harness the distributed nature of the data across clients.

FedProx[35]: By incorporating this proximal term, the algorithm ensures that the local updates from clients do not deviate significantly from the current global model. This design not only enhances the stability and convergence of the learning process but also ensures robustness in heterogeneous and asynchronous federated environments.

4.4. Comparision Analysis and Convergence Analysis

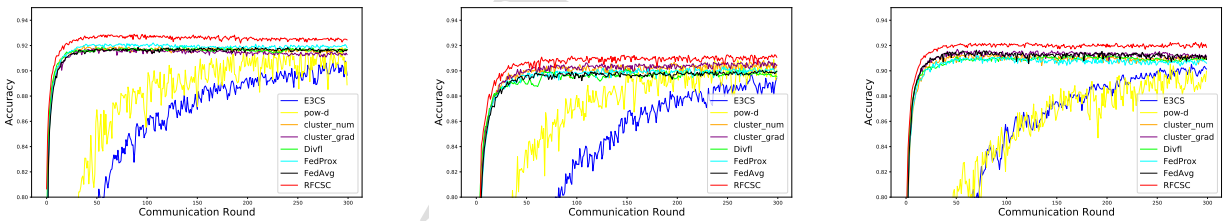
We show detailed comparison results, analyze the parameters, and offer further evaluations one after the other.

Based on the accuracy curves in Figures 2, 3, and 4, it can be observed that RFCSC exhibits excellent convergence across different datasets and under varying data skew conditions ($\alpha = 0.7, 0.8, 0.9$). In the FEMNIST dataset, RFCSC shows the fastest convergence speed and the highest convergence accuracy for all α values, particularly demonstrating strong robustness under high data skew conditions ($\alpha = 0.9$). In the more complex CIFAR-10 dataset, RFCSC converges to a stable accuracy after approximately 100 communication rounds, while other methods, such as E3CS and Pow-d, converge more slowly and exhibit less stable convergence processes. For the VIR dataset, RFCSC consistently shows high accuracy and stability under all conditions, especially in high data skew conditions ($\alpha = 0.9$). Although the accuracy curves for all methods show increased fluctuation, RFCSC still maintains a high final accuracy. These convergence analyses indicate that RFCSC has significant advantages in handling non-independent and identically distributed data and complex data distributions, further validating its effectiveness and robustness in practical applications.

From Tab.1, we can observe that RFCSC achieves higher test accuracy on all three datasets compared to other methods. The results on FEMNIST are shown in Fig.2. In our experimental setup, we opted for a client selection rate of 40% and investigated data skewness levels set at 0.7, 0.8 and 0.9, spanning over 300 communication rounds. In the figures, we have omitted the results for the IID data

Table 1
Comparison test accuracy on the three datasets

Dataset	β	α	E3CS	Pow-d	DivFL	Cl-num	Cl-grad	FedProx	FedAvg	RFCSC
Femnist	0.2	IID	93.19	93.17	93.23	93.01	92.91	92.98	93.28	**93.36**
		0.7	87.25	92.19	92.09	92.14	92.07	92.14	91.75	**92.53**
		0.8	88.57	90.96	91.35	91.78	91.66	91.75	91.41	**92.17**
		0.9	85.52	90.52	89.91	90.46	89.49	90.57	91.11	**91.31**
	0.4	IID	93.24	92.97	92.87	93.06	93.03	92.95	93.06	**93.32**
		0.7	90.36	91.62	91.97	92.01	91.81	92.21	91.86	**92.89**
		0.8	90.52	90.53	91.24	91.39	91.74	91.32	91.64	**92.25**
		0.9	89.69	90.96	90.13	90.83	90.84	90.31	90.03	**91.27**
Cifar-10	0.2	IID	57.63	54.51	57.99	58.26	59.12	56.33	59.27	**60.05**
		0.7	54.66	56.73	52.91	53.48	56.05	55.94	56.82	**59.51**
		0.8	47.17	54.41	51.92	51.53	55.26	54.08	55.19	**58.21**
		0.9	38.88	51.41	47.74	50.03	52.72	48.77	49.65	**55.15**
	0.4	IID	54.65	56.69	56.31	53.08	57.31	55.09	56.19	**61.41**
		0.7	56.64	56.49	55.03	55.06	55.29	54.61	56.01	**59.05**
		0.8	46.22	53.42	53.17	53.95	54.01	53.38	53.93	**57.43**
		0.9	45.91	53.12	51.13	51.18	52.04	51.72	52.09	**55.38**
VIR	0.2	IID	70.84	68.92	72.13	72.21	71.43	73.26	74.35	**75.64**
		0.7	76.99	76.97	77.19	76.55	77.09	74.97	77.14	**78.82**
		0.8	76.55	74.75	76.84	72.44	76.47	76.67	76.77	**77.61**
		0.9	73.22	72.76	72.56	74.01	72.04	73.12	73.08	**77.39**
	0.4	IID	70.56	67.13	72.97	72.71	71.15	73.05	74.65	**75.91**
		0.7	77.09	76.18	76.23	76.61	75.23	75.98	75.01	**78.05**
		0.8	77.26	77.21	75.64	77.44	75.88	75.56	76.57	**77.47**
		0.9	74.95	74.68	74.51	74.11	74.13	74.45	75.86	**76.95**

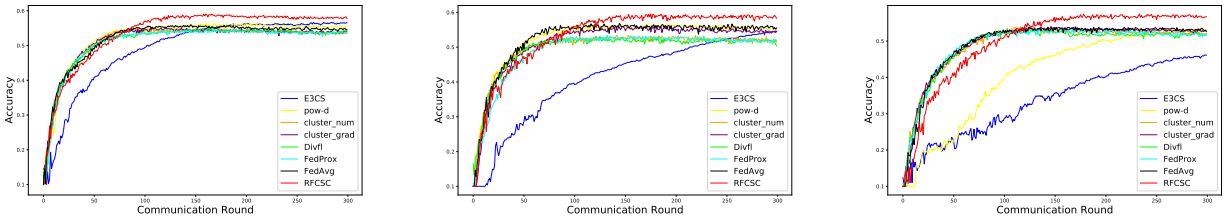


(a) $\beta = 0.4$ $\alpha = 0.7$ (b) $\beta = 0.4$ $\alpha = 0.8$ (c) $\beta = 0.4$ $\alpha = 0.9$

Figure 2: Accuracy curves on FEMNIST

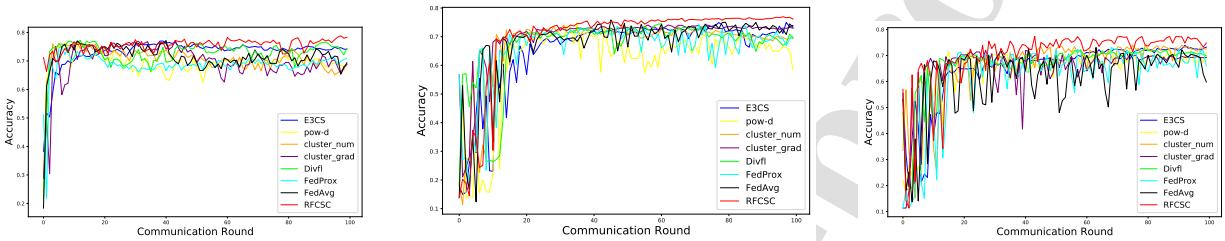
scenario, given that all methods consistently exhibit satisfactory performance in this context. Our observations indicate that irrespective of the variations in client data skewness, RFCSC consistently outperforms in terms of accuracy and convergence rate. Both E3CS and pow-d manifested suboptimal outcomes in comparison. The performance metrics for other methodologies were relatively analogous. Intriguingly, FedAvg, with its stochastic client selection approach, exhibited superior performance at an alpha value of 0.9. This potentially suggests that deterministic strategies employed by other methods might be susceptible to local optima. The superior performance of RFCSC over FedAvg further underscores the efficacy of our proposed approach.

For CIFAR-10, the experimental configuration mirrors that of FEMNIST. As discerned from the Tab.1, in the IID data context, all methodologies exhibit commendable outcomes. Outside the IID paradigm, irrespective of the client selection ratio, RFCSC consistently outperforms its counterparts by a margin of at least 3%. Remarkably, even under the most pronounced data skewness of 0.9, it guarantees an accuracy threshold exceeding 55%. As shown in Fig.3, RFCSC outperforms other methods on the CIFAR-10 dataset; however, it still has some drawbacks. RFCSC exhibits a relatively slow initial convergence speed, resulting in a slower accuracy improvement during the early communication rounds (approximately the first 50 rounds). Beyond



(a) $\beta = 0.4$ $\alpha = 0.7$ (b) $\beta = 0.4$ $\alpha = 0.8$ (c) $\beta = 0.4$ $\alpha = 0.9$

Figure 3: Accuracy curves on CIFAR-10



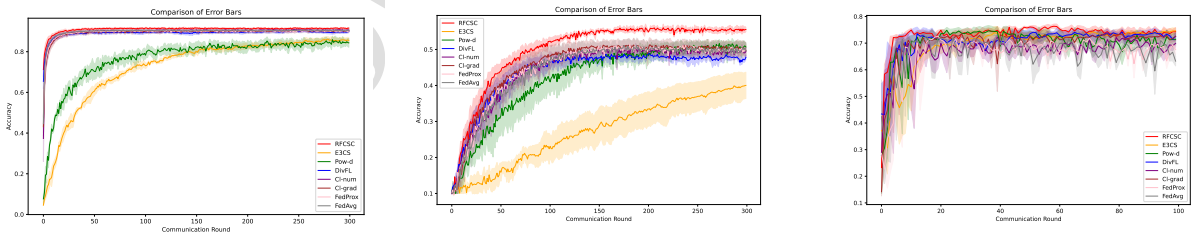
(a) $\beta = 0.4$ $\alpha = 0.7$ (b) $\beta = 0.4$ $\alpha = 0.8$ (c) $\beta = 0.4$ $\alpha = 0.9$

Figure 4: Accuracy curves on VIR

100 aggregation iterations, the superiority of RFCSC becomes palpable. This can be attributed to the inherent complexity of CIFAR-10 relative to FEMNIST, coupled with the preliminary experience accumulation phase intrinsic to reinforcement learning, which momentarily decelerates convergence.

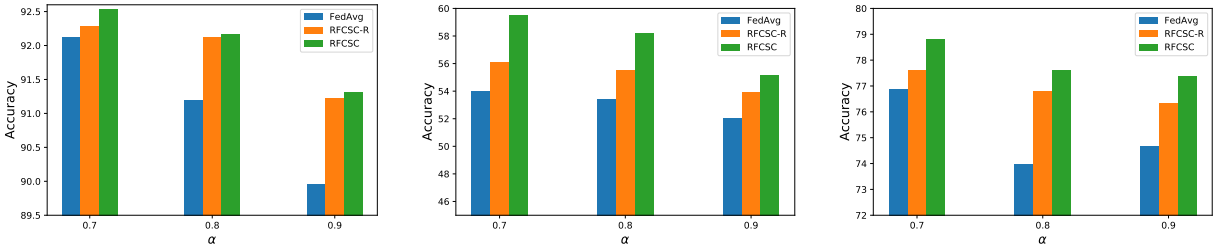
The results on VIR are shown in Tab.1 and Fig.4. From the presented tables and figure representations, it is discernible that all methodologies exhibit enhanced performance under Non-IID conditions as opposed to IID scenarios. Notably, all strategies manifest pronounced oscillations in their performance metrics. This phenomenon can be ascribed to the inherent complexities of public safety images, which are susceptible to a myriad of external perturbations, encompassing variations in lighting conditions, meteorological influences, and pixel resolutions. Such intricacies introduce considerable noise and outliers in the imagery data. Within an IID context, these perturbations are likely to be uniformly dispersed, thereby posing substantial challenges for model adaptability and accuracy. Conversely, in Non-IID settings, the models might exhibit a heightened resilience to such disturbances, facilitating more adept handling of the noise and outliers. Consistently, RFCSC manifests optimal performance metrics on the VIR dataset.

Fig.5 presents the error bars for all methods on the FEMNIST, CIFAR-10, and VIR datasets. We observed that the error bars for RFCSC are relatively narrow across all datasets, indicating a low variance in its predictions. This indicates that the model's performance remains consistent and reliable across different communication rounds. This stability is particularly evident on the FEMNIST and CIFAR-10 datasets, where the error bars remain within a narrow range even as accuracy steadily increases. Even on the more complex VIR dataset, the error bars for RFCSC remain relatively narrow, further highlighting the algorithm's stability and consistency when handling different types of data.



(a) FEMNIST (b) CIFAR-10 (c) VIR

Figure 5: Accuracy Curves with Error Bars on Different Datasets



(a) FEMNIST (b) CIFAR-10 (c) VIR

Figure 6: Ablation results to analyze the effect of the two components

4.5. Ablation Analysis

In this section, we examine how each component of our proposed model contributes to its overall performance. Compared to the traditional FedAvg method, our approach includes two key improvements: our approach adds reinforcement learning for dynamic client selection and uses adaptive model gradient compression mechanism. The client selection strategy with reinforcement learning is called RFCSC-R. When we add adaptive model gradient compression mechanism to this strategy, we get the complete RFCSC framework. We compare both RFCSC and RFCSC-R with FedAvg to show the importance of each component. In our evaluation, we compared our methods using three different datasets, keeping client selection rate as 0.2. The experimental results in Fig. 6 show that RFCSC-R (orange bars) achieves a significant performance improvement over the traditional FedAvg (blue bars) method across all three datasets (FEMNIST, CIFAR-10, and VIR). The RFCSC-R method effectively selects clients that contribute more to the global model by incorporating a reinforcement learning client selection strategy based on data prototypes. This approach enhances model accuracy under different levels of non-independent and Non-IID conditions(α values). By leveraging data prototypes and reinforcement learning strategies, RFCSC-R can better select appropriate clients under various data distributions and conditions, resulting in higher model accuracy compared to the traditional FedAvg method. Moreover, incorporating adaptive model gradient compression mechanism within RFCSC further enhances accuracy.

4.6. Communication Cost

In the evaluation of communication cost, for each dataset, we designated a specific target accuracy threshold. We recorded the number of federated rounds required for each method to reach the target accuracy, along with the memory size of the gradients. The slash indicates that the method did not achieve the target accuracy. In Tab.2, an exhaustive evaluation of communication cost among various federated learning strategies is delineated. Our proposed method, RFCSC, exhibits advantages across all evaluated datasets. Specifically, on the Femnist dataset, RFCSC realizes an accuracy of 91.5% with only 13 communication rounds, a significant reduction compared to other methods. In the more complex CIFAR-10 dataset, RFCSC exhibits a noticeable advantage in both accuracy and the number of communication rounds, slightly outperforming FedAvg. Furthermore, while the size of uploaded parameters for RFCSC fluctuates across datasets, it consistently manifests a smaller parameter size and fewer communication rounds on certain datasets, such as VIR and CIFAR-10. This underscores RFCSC's heightened efficiency in communication. Collectively, these results provide compelling evidence of RFCSC's efficacy and adaptability in federated learning, particularly in environments where bandwidth is at a premium or communication costs are a concern.

4.7. Parameter Sensitivity Analysis

4.7.1. The impact of different compression rates on model performance

Figure 7 illustrates the impact of different compression rates on model accuracy across three datasets (FEMNIST, CIFAR-10, VIR). In the FEMNIST dataset, the RFCSC method maintains the highest accuracy across all compression rates, showing its stability and robustness on simpler datasets. For the more complex CIFAR-10 dataset, RFCSC achieves the highest accuracy at a compression rate of 0.4, indicating that the strategy effectively balances model performance and communication cost when dealing with complex data. On the VIR dataset, RFCSC performs best at a compression rate of 0.5, maintaining excellent performance even when handling complex and noisy data. From the Fig. 7(b) and Fig. 7(c), it is evident that, in the more complex CIFAR-10 and VIR datasets, RFCSC achieves its highest accuracy at compression rates of 0.4 and 0.5, respectively. These optimal compression rates are found by adaptive compression mechanism. It means that a suitable compression rate can improve model performance. From the Fig. 7(a), we can find that in the relatively simpler FEMNIST dataset, the impact of compression rate on RFCSC is minimal, while other methods exhibit significant fluctuations. These experimental results support the effectiveness of the adaptive compression strategy, demonstrating that RFCSC can maintain high model accuracy by dynamically adjusting the compression rate across different datasets and compression

Table 2
Communication Cost

Method	Femnist Target Acc=91.5%		Cifar-10 Target Acc=55%		VIR Target Acc=75%	
	rounds	param(MB)	rounds	param(MB)	rounds	param(MB)
E3CS	226	219.61	162	6.02	11	155.21
Pow-d	25	24.32	102	3.78	9	126.99
DivFL	26	25.34	163	6.06	10	141.11
Cluster_num	25	24.33	108	4.01	16	225.76
Cluster_grad	29	28.28	120	4.46	96	1354.61
FedProx	23	21.51	\	\	12	169.32
FedAvg	32	31.16	89	3.31	11	155.21
RFCSC	13	12.61	87	3.23	6	84.66

Table 3
Test Accuracy for different Compression Rates

Dataset	σ	E3CS	Pow-d	DivFL	Cl-num	Cl-grad	FedProx	FedAvg	RFCSC
Femnist	0	84.45	84.26	89.88	90.43	90.65	90.51	89.96	**91.67**
	0.2	86.23	84.84	90.09	90.94	90.73	90.11	90.17	**91.53**
	0.4	85.51	83.92	89.91	90.45	89.49	90.57	91.11	**91.31**
	0.6	86.12	87.47	89.94	90.47	90.68	90.37	90.38	**91.29**
	0.8	84.41	85.69	89.17	89.51	90.35	90.15	90.11	**91.18**
Cifar10	0	34.53	50.86	48.25	52.69	50.07	51.65	52.07	**53.93**
	0.2	32.85	51.28	48.55	49.64	50.02	50.21	50.34	**53.65**
	0.4	38.82	51.41	47.74	50.03	52.72	48.77	49.65	**55.15**
	0.6	36.62	49.44	48.49	47.48	50.62	50.79	49.15	**52.47**
	0.8	35.16	49.62	46.61	48.06	48.21	46.77	47.48	**51.21**
VIR	0	72.53	75.27	73.11	75.83	74.26	75.54	74.68	**76.33**
	0.2	72.55	75.02	73.27	74.92	71.96	75.88	71.42	**76.11**
	0.5	73.22	72.76	72.56	74.01	72.04	73.12	73.08	**77.39**
	0.8	72.43	74.38	72.65	73.32	72.46	70.36	73.05	**75.66**

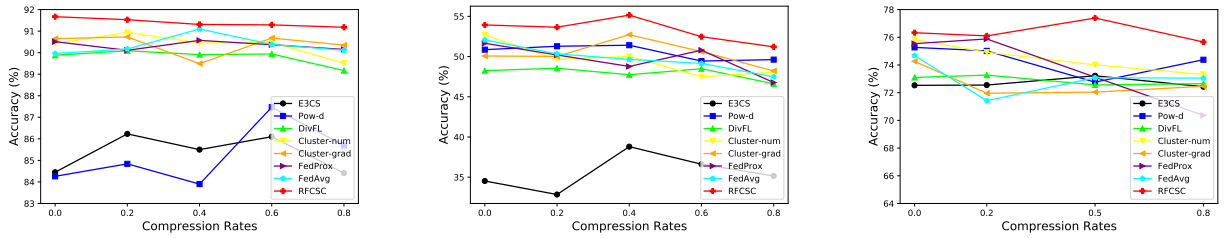
levels, highlighting its practical value and advantages in federated learning.

4.7.2. The impact of η_1 and η_2 on model performance

In this section, we analysis parameter sensitivity. In this experiment, We select the representative CIFAR-10 datase. In Fig. 8 (a)-(c) present the results of parameter sensitivity analysis. This section predominantly examines the impact of η_1 and η_2. η_1 is the client's local learning rate and η_2 is the learning rate updating the reinforcement learning online network. For learning rate η_1 ranging between 0.001 and 0.01, and learning rate η_2 ranging from 0.001 to 0.00001. From Fig. 8(a), we can find that the best test accuracy is at (0.001,0.00005). The learning rate of local client models exerts a significant impact on the overall model, whereas the influence of the learning rate pertaining to the online reinforcement learning model on the overall model is comparatively marginal.

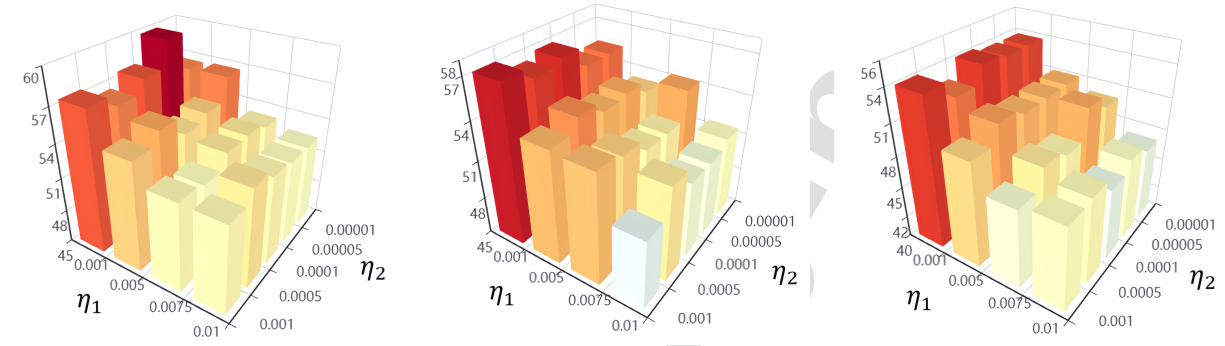
5. Conclusion

To address the Non-IID data distribution across various clients in federated learning and to reduce the communication cost inherent in federated learning, we propose a Reinforcement Federated Client Selection and Compression method, which combines reinforcement learning with adaptive compression mechanism. We devise a reinforcement learning client selection strategy with data prototypes as the state space. By using these data prototypes, we can perceive the primary features of client network parameters and data patterns, thereby selecting optimal clients for dynamic aggregation. Furthermore, by using adaptively compressing the model of local clients, valuable gradient information is preserved and effectively reduce communication cost. RFCSC facilitates collaborative training across multiple parties. Extensive experiments on various datasets have validated the effectiveness of the proposed method.



(a) FEMNIST (b) CIFAR-10 (c) VIR

Figure 7: Analysis of the influence of the compression rate



(a) $\alpha = 0.7$ (b) $\alpha = 0.8$ (c) $\alpha = 0.9$

Figure 8: Parameter sensitivity analysis of learning rates

Acknowledgment

This work was supported by the National Natural Science Foundation of China (U22B2038, 62172056, 62192784), the 8th Young Elite Scientists Sponsorship Program by CAST (2022QNRC001) and the Postgraduate Education and Teaching Reform Project of Beijing University of Posts and Telecommunications(2023Y028).

CRediT authorship contribution statement

Zhenhui Pan: Methodology, Software, Writing - original draft. **Yawen Li:** Writing - review & editing. **Zeli Guan:** Writing - review & editing. **Meiyu Liang:** Writing - review & editing. **Ang Li:** Writing - review & editing. **Jia Wang:** Writing - review & editing. **Feifei Kou:** Software & editing

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Şeymanur Aktı, Ferda Ofli, Muhammad Imran, and Hazim Kemal Ekenel. Fight detection from still images in the wild. In *Proceedings of the IEEE/CVF winter conference on applications of computer vision*, pages 550–559, 2022.

[2] Mauricio Perez, Alex C Kot, and Anderson Rocha. Detection of real-world fights in surveillance videos. In *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 2662–2666, 2019.

[3] Vishwanath A Sindagi and Vishal M Patel. A survey of recent advances in cnn-based single image crowd counting and density estimation. *Pattern Recognition Letters*, 107:3–16, 2018.

[4] Yu Han Liu. Feature extraction and image recognition with convolutional neural networks. In *Journal of Physics: Conference Series*, volume 1087, page 062032. IOP Publishing, 2018.

[5] Swathikiran Sudhakaran and Oswald Lanz. Learning to detect violent videos using convolutional long short-term memory. In *2017 14th IEEE international conference on advanced video and signal based surveillance (AVSS)*, pages 1–6, 2017.

[6] Tian Li, Anit Kumar Sahu, Ameet Talwalkar, and Virginia Smith. Federated learning: Challenges, methods, and future directions. *IEEE signal processing magazine*, 37(3):50–60, 2020.

[7] SungHwan Moon and Won Hee Lee. Privacy-preserving federated learning in healthcare. In *2023 International Conference on Electronics, Information, and Communication (ICEIC)*, pages 1–4, 2023.

[8] WANG Luping, WANG Wei, and LI Bo. Cmfl: Mitigating communication overhead for federated learning. In *2019 IEEE 39th international conference on distributed computing systems (ICDCS)*, pages 954–964, 2019.

[9] Enrico Russo, Maurizio Palesi, Salvatore Monteleone, Davide Patti, Andrea Mineo, Giuseppe Ascia, and Vincenzo Catania. Dnn model compression for iot domain-specific hardware accelerators. *IEEE Internet of Things Journal*, 9(9):6650–6662, 2021.

[10] Yu Cheng, Duo Wang, Pan Zhou, and Tao Zhang. Model compression and acceleration for deep neural networks: The principles, progress, and challenges. *IEEE Signal Processing Magazine*, 35(1):126–136, 2018.

[11] Zhuang Liu, Jianguo Li, Zhiqiang Shen, Gao Huang, Shoumeng Yan, and Changshui Zhang. Learning efficient convolutional networks through network slimming. In *Proceedings of the IEEE international conference on computer vision*, pages 2736–2744, 2017.

[12] Chenglong Zhao, Bingbing Ni, Jian Zhang, Qiwei Zhao, Wenjun Zhang, and Qi Tian. Variational convolutional neural network pruning. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 2780–2789, 2019.

[13] Prateeth Nayak, David Zhang, and Sek Chai. Bit efficient quantization for deep neural networks. In *2019 Fifth Workshop on Energy Efficient Machine Learning and Cognitive Computing-NeurIPS Edition (EMC2-NIPS)*, pages 52–56, 2019.

[14] Tara N Sainath, Brian Kingsbury, Vikas Sindhwani, Ebru Arisoy, and Bhuvana Ramabhadran. Low-rank matrix factorization for deep neural network training with high-dimensional output targets. In *2013 IEEE international conference on acoustics, speech and signal processing*, pages 6655–6659, 2013.

[15] Jang Hyun Cho and Bharath Hariharan. On the efficacy of knowledge distillation. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 4794–4802, 2019.

[16] Tao Lu, Qing Liu, Xubin He, Huizhang Luo, Eric Suchyta, Jong Choi, Norbert Podhorszki, Scott Klasky, Mathew Wolf, Tong Liu, et al. Understanding and modeling lossy compression schemes on hpc scientific data. In *2018 IEEE International Parallel and Distributed Processing Symposium (IPDPS)*, pages 348–357, 2018.

[17] Sindhu Padakandla. A survey of reinforcement learning algorithms for dynamically varying environments. *ACM Computing Surveys (CSUR)*, 54(6):1–25, 2021.

[18] Jens Kober, J Andrew Bagnell, and Jan Peters. Reinforcement learning in robotics: A survey. *The International Journal of Robotics Research*, 32(11):1238–1274, 2013.

[19] Ahmad El Sallab, Mohammed Abdou, Etienne Perot, and Senthil Yogamani. Deep reinforcement learning framework for autonomous driv-ing. *stat*, 1050:8, 2017.

[20] Qingyun Sun, Jianxin Li, Hao Peng, Jia Wu, Yuanxing Ning, Philip S Yu, and Lifang He. Sugar: Subgraph neural network with reinforcement pooling and self-supervised mutual information mechanism. In *Proceedings of the Web Conference 2021*, pages 2081–2091, 2021.

[21] Decheng Liu, Zhan Dang, Chunlei Peng, Yu Zheng, Shuang Li, Nannan Wang, and Xinbo Gao. Fedforgery: generalized face forgery detection with residual federated learning. *IEEE Transactions on Information Forensics and Security*, 2023.

[22] Nehemia Sugianto, Dian Tjondronegoro, Rosemary Stockdale, and Elizabeth Irenne Yuwono. Privacy-preserving ai-enabled video surveillance for social distancing: Responsible design and deployment for public spaces. *Information Technology & People*, 37(2):998–1022, 2024.

[23] David G Lowe. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60:91–110, 2004.

[24] Svetlana Lazebnik, Cordelia Schmid, and Jean Ponce. Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories. In *2006 IEEE computer society conference on computer vision and pattern recognition (CVPR'06)*, volume 2, pages 2169–2178, 2006.

[25] Navneet Dalal and Bill Triggs. Histograms of oriented gradients for human detection. In *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)*, volume 1, pages 886–893, 2005.

[26] Jeho Nam, Masoud Alghoniemy, and Ahmed H Tewfik. Audio-visual content-based violent scene characterization. In *Proceedings 1998 International Conference on Image Processing. ICIP98 (Cat. No. 98CB36269)*, volume 1, pages 353–357, 1998.

[27] Nianyin Zeng, Xinyu Li, Peishu Wu, Han Li, and Xin Luo. A novel tensor decomposition-based efficient detector for low-altitude aerial objects with knowledge distillation scheme. *IEEE/CAA Journal of Automatica Sinica*, 11(2):487–501, 2024.

[28] Peishu Wu, Zidong Wang, Han Li, and Nianyin Zeng. Kd-par: A knowledge distillation-based pedestrian attribute recognition model with multi-label mixed feature learning network. *Expert Systems with Applications*, 237:121305, 2024.

[29] Han Li, Zidong Wang, Chengbo Lan, Peishu Wu, and Nianyin Zeng. A novel dynamic multiobjective optimization algorithm with non-inductive transfer learning based on multi-strategy adaptive selection. *IEEE Transactions on Neural Networks and Learning Systems*, 2023.

[30] Prithvijit Chattopadhyay, Yogesh Balaji, and Judy Hoffman. Learning to balance specificity and invariance for in and out of domain generalization. In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part IX 16*, pages 301–318. Springer, 2020.

[31] Yingjun Du, Jun Xu, Huan Xiong, Qiang Qiu, Xiantong Zhen, Cees GM Snoek, and Ling Shao. Learning to learn with variational information bottleneck for domain generalization. pages 200–216, 2020.

[32] Fengchun Qiao, Long Zhao, and Xi Peng. Learning to learn single domain generalization. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 12556–12565, 2020.

[33] Peishu Wu, Zidong Wang, Baixun Zheng, Han Li, Fuad E Alsaadi, and Nianyin Zeng. Aggn: Attention-based glioma grading network with multi-scale feature extraction and multi-modal information fusion. *Computers in biology and medicine*, 152:106457, 2023.

[34] Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, and Blaise Aguera y Arcas. Communication-efficient learning of deep networks from decentralized data. In *Artificial intelligence and statistics*, pages 1273–1282, 2017.

[35] Tian Li, Anit Kumar Sahu, Manzil Zaheer, Maziar Sanjabi, Ameet Talwalkar, and Virginia Smith. Federated optimization in heterogeneous networks. *Proceedings of Machine learning and systems*, 2:429–450, 2020.

[36] Di Wu, Rehmat Ullah, Paul Harvey, Peter Kilpatrick, Ivor Spence, and Blesson Varghese. Fedadapt: Adaptive offloading for iot devices in federated learning. *IEEE Internet of Things Journal*, 9(21):20889–20901, 2022.

[37] Hongyi Wang, Mikhail Yurochkin, Yuekai Sun, Dimitris Papailiopoulos, and Yasaman Khazaeni. Federated learning with matched averaging. In *International Conference on Learning Representations*, 2019.

[38] Tzu-Ming Harry Hsu, Hang Qi, and Matthew Brown. Measuring the effects of non-identical data distribution for federated visual classification. *arXiv preprint arXiv:1909.06335*, 2019.

[39] Ron Banner, Itay Hubara, Elad Hoffer, and Daniel Soudry. Scalable methods for 8-bit training of neural networks. *Advances in neural information processing systems*, 31, 2018.

[40] Yawen Li, Wenling Li, and Zhe Xue. Federated learning with stochastic quantization. *International Journal of Intelligent Systems*, 37(12):11600–11621, 2022.

[41] Elliot J Crowley, Jack Turner, Amos Storkey, and Michael O'Boyle. A closer look at structured pruning for neural network compression. *arXiv preprint arXiv:1810.04622*, 2018.

[42] Hao Li, Asim Kadav, Igor Durdanovic, Hanan Samet, and Hans Peter Graf. Pruning filters for efficient convnets. *arXiv preprint arXiv:1608.08710*, 2016.

[43] Lucas Theis, Iryna Korshunova, Alykhan Tejani, and Ferenc Huszár. Faster gaze prediction with dense networks and fisher pruning. *arXiv preprint arXiv:1801.05787*, 2018.

[44] Hao Wang, Zakhary Kaplan, Di Niu, and Baochun Li. Optimizing federated learning on non-iid data with reinforcement learning. In *IEEE INFOCOM 2020-IEEE Conference on Computer Communications*, pages 1698–1707, 2020.

[45] Christopher JCH Watkins and Peter Dayan. Q-learning. *Machine learning*, 8:279–292, 1992.

[46] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Alex Graves, Ioannis Antonoglou, Daan Wierstra, and Martin Riedmiller. Playing atari with deep reinforcement learning. *arXiv preprint*

arXiv:1312.5602, 2013.

[47] Zeli Guan, Yawen Li, Zhenhui Pan, Yuxin Liu, and Zhe Xue. Rfdg: Reinforcement federated domain generalization. *IEEE Transactions on Knowledge and Data Engineering*, 2023.

[48] Yue Tan, Guodong Long, Lu Liu, Tianyi Zhou, Qinghua Lu, Jing Jiang, and Chengqi Zhang. Fedproto: Federated prototype learning across heterogeneous clients. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pages 8432–8440, 2022.

[49] Shiliang Sun, Yuhan Liu, and Liang Mao. Multi-view learning for visual violence recognition with maximum entropy discrimination and deep features. *Information Fusion*, 50:43–53, 2019.

[50] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6):84–90, 2017.

[51] Sebastian Caldas, Sai Meher Karthik Duddu, Peter Wu, Tian Li, Jakub Konečnỳ, H Brendan McMahan, Virginia Smith, and Ameet Talwalkar. Leaf: A benchmark for federated settings. *arXiv preprint arXiv:1812.01097*, 2018.

[52] Qinbin Li, Bingsheng He, and Dawn Song. Model-contrastive federated learning. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 10713–10722, 2021.

[53] Tiansheng Huang, Weiwei Lin, Li Shen, Keqin Li, and Albert Y Zomaya. Stochastic client selection for federated learning with volatile clients. *IEEE Internet of Things Journal*, 9(20):20055–20070, 2022.

[54] Yae Jee Cho, Jianyu Wang, and Gauri Joshi. Client selection in federated learning: Convergence analysis and power-of-choice selection strategies. *arXiv preprint arXiv:2010.01243*, 2020.

[55] Yann Fraboni, Richard Vidal, Laetitia Kameni, and Marco Lorenzi. Clustered sampling: Low-variance and improved representativity for clients selection in federated learning. In *International Conference on Machine Learning*, pages 3407–3416, 2021.

[56] Ravikumar Balakrishnan, Tian Li, Tianyi Zhou, Nageen Himayat, Virginia Smith, and Jeff Bilmes. Diverse client selection for federated learning via submodular maximization. In *International Conference on Learning Representations*, 2022.

Zhenhui Pan is a Ph.D. candidate in Computer Science and Technology at the Beijing University of Posts and Telecommunications. He received a B.S. degree from University of Liverpool in 2018 and an M.S. degree from University College London in 2019, all related to Statistics science. His research interests include federated learning and machine learning.

Yawen Li is an associate professor at the School of Economics and Management, Beijing University of Posts and Telecommunications. She received her Ph.D. in Innovation, Entrepreneurship, and Strategy from Tsinghua University in 2018. Her research interest focuses on artificial intelligence, collaborative innovation, the development of science parks, and the scientific productivity of firms.

Zeli Guan is a Ph.D. candidate of School of Computer Science, Beijing University of Posts and Telecommunications. His mainly research directions for federated Learning, graph neural network and machine learning. His research papers have been published in or accepted by International Journal of Intelligent Systems, Computational Intelligence and Neuroscience, and Visual Informatics

Meiyu Liang received the Ph.D. degree in Computer Science and technology from Beijing University of Posts and Telecommunications, Beijing, China, in 2014. She ever did postdoctoral research in School of Computer Science from Beijing University of Posts and Telecommunications from 2014 to 2016. She is currently a professor and doctoral supervisor in School of Computer Science, Beijing University of Posts and Telecommunications, Beijing, China. Her research interests include artificial intelligence, machine learning, multi-modal information processing and mining, cross-media semantic learning and search.

Ang Li is a Ph.D. candidate in Computer Science and Technology at the Beijing University of Posts and Telecommunications. He received a B.S. degree from the Nanchang Hangkong University in 2015 and an M.S. degree from the Beijing University of Posts and Telecommunications in 2019, all related to computer science. His research interests include information retrieval, scholar profiling, and data mining.

Jia Wang is a Ph.D. candidate in Computer Science and Technology at the Beijing University of Posts and Telecommunications. He received a B.S. degree from the Beijing University of Posts and Telecommunications

in 2021, all related to computer science. His research interests include federated learning and data mining.

Feifei Kou is a lecturer at the School of Computer Science (National Pilot School of Software Engineering), Beijing University of Posts and Telecommunications, China. She received her Ph.D. degree at the School of Computer Science from Beijing University of Posts and Telecommunications in 2019. She ever did postdoctoral research in School of Computer Science (National Pilot School of Software Engineering), Beijing University of Posts and Telecommunications from 2019 to 2021. Her research interests include semantic learning, and multimedia information processing.

Declaration of interests

☐ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☒ The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Zhenhui Pan reports financial support was provided by National Natural Science Foundation of China. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.