|  |
| --- |
| *CR-Form-v12.2* |
| **PSEUDO CHANGE REQUEST** |
|  |
|  |  | **CR** |  | **rev** |  | **Current version:** | **0** |  |
|  |
| *For* [***HE******LP***](http://www.3gpp.org/3G_Specs/CRs.htm#_blank)*on using this form: comprehensive instructions can be found at* [*http://www.3gpp.org/Change-Requests*](http://www.3gpp.org/Change-Requests)*.* |
|  |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ***Proposed change affects:*** | UICC apps |  | ME | **X** | Radio Access Network |  | Core Network | **X** |

|  |
| --- |
|  |
| ***Title:***  | pCR on Federated learning |
|  |  |
| ***Source to WG:*** | Tencent |
| ***Source to TSG:*** | S4 |
|  |  |
| ***Work item code:*** | FS\_AI4Media |  | ***Date:*** | 07 |
|  |  |  |  |  |
| ***Category:*** |  |  | ***Release:*** |  |
|  | *Use one of the following categories:****F*** *(correction)****A*** *(mirror corresponding to a change in an earlier release)****B*** *(addition of feature),* ***C*** *(functional modification of feature)****D*** *(editorial modification)*Detailed explanations of the above categories canbe found in 3GPP [TR 21.900](http://www.3gpp.org/ftp/Specs/html-info/21900.htm). | *Use one of the following releases:Rel-8 (Release 8)Rel-9 (Release 9)Rel-10 (Release 10)Rel-11 (Release 11)…Rel-16 (Release 16)Rel-17 (Release 17)Rel-18 (Release 18)Rel-19 (Release 19)* |
|  |  |
| ***Reason for change:*** | Updates on federated learning workflow and corresponding control messages |
|  |  |
| ***Summary of change:*** |  |
|  |  |
| ***Consequences if not approved:*** |  |
|  |  |
| ***Clauses affected:*** | 5.2.4.2, 6.5 |
|  |  |
|  | **Y** | **N** |  |  |
| ***Other specs*** |  | **X** |  Other core specifications  | TS/TR ... CR ...  |
| ***affected:*** |  | **X** |  Test specifications | TS/TR ... CR ...  |
| ***(show related CRs)*** |  | **X** |  O&M Specifications | TS/TR ... CR ...  |
|  |  |
| ***Other comments:*** |  |
|  |  |
| ***This CR's revision history:*** |  |

Start of change 1

#### 5.2.4.2 Basic workflows

Figure 5.2.4-2 shows a basic workflow for distributed/federated learning with training in the UE, the results of which are aggregated in the network. Steps for the procedures shown are described below.



Figure 5.2.4-2: Basic workflow for distributed/federated learning between a UE and the network

During the initialization and establishment step, it is assumed that information related to the required features and detailed configurations are exchanged and negotiated between the network and UE. Information may include those related to UE device and network capabilities, AI/ML service information (e.g. service requirements, AI/ML model descriptions), and delivery methods. Such information may be used for the selection of a suitable partially trained AI/ML model for the service.

1. The UE Application and Network Application communicate to trigger distributed/federated learning, using the information from the initialization and establishment step.

2. A partially trained AI model is selected between the UE Application and Network Application.

3. The Network Application identifies the selected partially trained AI model in the AI model Repository/Provider.

4. The Federated Learning Engine optionally announces the eligibility criteria for participating in the federated evaluation/learning to the device. The criteria could contain various information such as the device's operating system, processor speed, available memory, characteristics of the data library, geographical location of the device, language setting, and other attributes.

5. The AI Model Access Function of an eligible device receives the partially trained AI model or its updated version

6. The Federated Learning Engine optionally announces the failure reporting criteria for the participating devices.

Option A: Model evaluation:

7. The Federated Learning Engine requests the UE to start the model evaluation. The evaluation mechanism and criteria are defined by the Federated learning Engine.

Note: Whether a user wants its device to participate in the evaluation, depends on the business agreement between the user and the network.

8. The Data Source passes the training input data to the AI model Training Engine.

9. The AI Model Training Engine performs the evaluation.

10. The evaluation results (or the failure messages, in the case of a failure) are delivered to the Federated Learning Engine.

11. Optionally, the device eligibility criteria may get updated depending on the evaluation results.

Option B: Federated training:

12. The Federated Learning Engine requests the UE to start the training.

Note: Whether a user wants its device to participate in the training, depends on the business agreement between the user and the network.

13. The Data Source passes the training input data to the AI model Training Engine.

14. The AI Model Training Engine performs the retraining of the model.

15. The updated model (or the failure messages, in the case of a failure) is delivered to the Federated Learning Engine.

16. The Federated Learning Engine performs training aggregation of training results from multiple UEs and updates the partially trained AI model.

17. The updated partially trained AI model is delivered to the UE as from step 5.

Note: As shown in the above call flow, the model evaluation and the federated learning may also occur in a sequence

End of change 1

Start of change 2

## 6.5 Metadata

[Editor’s note: Metadata may include metadata to describe AI/ML model types, metadata for split operation configurations, AI/ML operation endpoint capability metadata etc.]

### 6.5.1 Distributed/Federated learning

[Ed Note: Check consistency with distributed systems]

#### 6.5.1.1 Control information

##### 6.5.1.1.2 General

This clause describes a set of possible control messages for managing the training process, synchronization the training rounds, and defining the selection criteria for participating devices, or monitoring the convergence of the training process, in federated learning.

#### 6.5.1.2 Synchronization information

##### 6.5.1.2.1 Definition

Synchronization messages may be used to ensure that all devices start the training process simultaneously and progress at the same pace. For example, the server may send a synchronization message to all UEs to start a new round of training.

##### 6.5.1.2.2 Behavior

The network application sends a synchronization message to all UE applications to start a new round of training at the same time as described in step 1 of figure 5.2.4-2. The message contains the round number and may also contain a timestamp indicating when the training round should begin.

##### 6.5.1.2.3 Parameters

The possible parameters are:

- The Round\_number indicates the training round in a model training.

- The Start\_time indicates the start time of the training.

- The Duration indicates the desirable duration of the training. This value just shows an indication of the desirable time for completing the training round.

#### 6.5.1.3 Device eligibility information

##### 6.4.1.3.1 Definition

Device eligibility messages may be used to define the criteria for selecting the devices that will participate in the training process. For example, the server may send a device eligibility message to all devices that belong to the defined group by the application.

##### 6.5.1.3.2 Behavior

The Federated learning engine sends a device eligibility message to the AI model training engine to select the devices that meet certain criteria defined by the application as described in step 4 of figure 5.2.4-2. Depending on the number of criteria met, the application assigns a group id to the device. For example, the criteria could contain information about the device's operating system, processor speed, available memory, available image library (number of images…), geographical location of the device, language setting, and other attributes.

##### 6.5.1.3.3 Parameters

The possible parameters are:

- The Group\_id is used to assign a new id for the devices that meet the eligibility criteria of this message. If the device is eligible, it uses this value as one of its group ids and from now on, it reacts to messages with the same group id.

- The Application\_group\_id, is assigned by the application on the device and if that value is equal to the value of this field, then the device is eligible.

- The Hardware, Location, and Language parameters define the hardware, location, and language eligibility criteria respectively for the device.

- The Data\_library\_id defines the data library an eligible device shall have.

Note: if more than one eligibility field exists, the device needs to meet all criteria to become eligible.

#### 6.5.1.4 Model evaluation information

##### 6.5.1.4.1 Definition

Model evaluation messages may be used to evaluate the performance of the global model for each device and make decisions about the training process. After running the learning phase, a device sends a model evaluation message to the server that measures the accuracy of the model. The server can then decide whether to continue training for another round or stop.

Alternatively, this message may be used by the server to request the device to perform an evaluation of a newly downloaded global model.

##### 6.5.1.4.2 Behavior

For Federated learning engine sends the model evaluation messageto the AI model training engine in the UE containing the metrics to be used for evaluation such as accuracy or precision as described in step 7 of figure 5.2.4-2.

##### 6.5.1.4.3 Parameters

The possible parameters are:

- The Round\_number shows the round after which the evaluation is performed.

- The Metric\_number shows the number of metrics included in this message body.

- The Metric is one or more of the Name-Value pairs showing the name of the metric and the corresponding value obtained in the evaluation.

#### 6.5.1.5 Model update information

##### 6.5.1.5.1 Definition

Model update messages may be used to update the model parameters on the devices after each round of training. For example, the server may send a model update message to all devices to update the global model with the new model parameters.

Model update messages may also be used to update the global model on the server with the new parameters updated by the local training on the device.

##### 6.5.1.5.2 Behavior

The server may send a model update message to all devices to update the AI/ML model with the new model parameters as described in step 5 of figure 5.2.4-2. The message contains the model id of the AI/ML model to be updated, the updated model parameters that the UE will use to train the model in the next round, and the new model id when the parameters are updated.

After running the training locally, each AI Model training Engine in the UEs may send a model update message to the server with the updated parameters as described in step 15 of figure 5.2.4-2. Together with the received model evaluation message, the server can decide if the global model needs to be updated or not. The model update message then only contains the model id of the AI/ML model used for local training and the updated parameters.

##### 6.5.1.5.3 Parameters

The possible parameters are:

- The Parameters includes the new model vector of values.

- The New\_model\_id is the id of the new model when the server sends the model to one or more devices.

#### 6.5.1.6 Failure reporting information

##### 6.5.1.6.1 Definition

Error messages may be used to handle unexpected errors or exceptions that may occur during the training process. For example, the server may send an error message to all devices to handle a device failure or network disruption.

##### 6.5.1.6.2 Behavior

The server sends a request to all devices to report a device failure or network disruption as described in step 6 of figure 5.2.4-2. For example, if a device fails to send its model parameters back to the server, the device should notify the server so that the device has been removed from the training process.

The AI Model training engine in the UE sends a failure message to the Federated learning engine in the server if a failure occurs as described in step 15 of figure 5.2.4-2.

##### 6.5.1.6.3 Parameters

The Message describes the reason for the failure.

End of change 2