**3GPP TSG-RAN WG3 Meeting #114-e *R3-21xxxx***

**11-11 Aug 2021, E-meeting**

**Title:** TP to 37.817 on AI/ML based network energy saving

**Source:** ZTE

**Agenda item:** 18.4.1

**Document Type:** Other

# 1. Introduction

This TP tries to reflect agreement on the solution of AI/ML-based network energy saving from CB # AIRAN2\_ESSolution.

# 5. Reference

1. R3-215909, SoD\_CB # AIRAN2\_ESSolution

# Annex – TP for TR 37.817

//////////////////////////////////////////////////////////////irrelevant operations skipped/////////////////////////////////////////////////////////////////////

## 5.1 Network Energy Saving

### 5.1.1 Use case description

//////////////////////////////////////////////////////////////irrelevant operations skipped/////////////////////////////////////////////////////////////////////

### 5.1.2 Solutions and standard impacts

*Editor Note: Capture the solutions for the use case, including potential standard impacts on existing Nodes, functions, and interfaces*

The following solutions can be considered for supporting AI/ML-based network energy saving:

* AI/ML Model Training is located in the OAM and AI/ML Model Inference is located in the gNB.
* AI/ML Model Training and AI/ML Model Inference are both located in the gNB.

Note: gNB is also allowed to continue model training based on AI/ML model trained in the OAM

In case of CU-DU split architecture, the following solutions are possible:

* AI/ML Model Training is located in the OAM and AI/ML Model Inference is located in the gNB-CU.

AI/ML Model Training and Model Inference are both located in the gNB-CU.

#### 5.1.2.1 Model Training at OAM and Model Inference at NG-RAN

In this solution, NG-RAN predicts energy saving decisions by AI/ML model trained from OAM.



Figure 5.1.2.1-1. Model Training at OAM, Model Inference at NG-RAN

Step 0: NG-RAN node 1 is assumed to have a AI/ML model trained by OAM, NG-RAN node 2 is assumed to have a AI/ML model trained by OAM optionally.

Step 1: NG-RAN node 2 sends the required input data to NG-RAN node 1 for model inference of AI/ML-based network energy saving.

Step 2: UE sends UE measurement report to NG-RAN node 1. (FFS on if triggered)

Step 3: Based on local inputs of NG-RAN node 1 and received inputs from NG-RAN node 2, NG-RAN node 1 generates model inference output(s) (e.g. energy saving strategy, handover strategy, etc).

Step 4: NG-RAN node 1 optionally selects the most appropriate target cell for each UE before it performs handover and goes to the predicted energy state, if needed.

Step 5: NG-RAN node 1 and NG-RAN node 2 provide feedback to OAM.

#### 5.1.2.2 Model Training and Model Inference at NG-RAN

In this solution, NG-RAN is responsible for model training and generates energy saving decisions.

Editor’s Notes: FFS on data collection.



Figure 5.1.2.2-1. Model Training and Model Inference at NG-RAN

Step 1: NG-RAN node 1 trains AI/ML model for AI/ML-based energy saving based on collected data. NG-RAN node 2 is assumed to have AI/ML model for AI/ML-based energy saving optionally, which can also generate predicted results/actions.

Step 2: NG-RAN node 2 sends the required input data to NG-RAN node 1 for model inference of AI/ML-based network energy saving.

Step 3: UE sends UE measurement report to NG-RAN node 1. (FFS on if triggered)

Step 4: Based on local inputs of NG-RAN node 1 and received inputs from NG-RAN node 2, NG-RAN node 1 generates model inference output (e.g. energy saving strategy, handover strategy, etc).

Step 5: NG-RAN node 1 optionally selects the most appropriate target cell for each UE before it performs handover and goes to the predicted energy state, if needed.

Step 6: NG-RAN node 2 provides feedback to NG-RAN node 1.

#### 5.1.2.3 Input of AI/ML-based Network Energy Saving

To predict the optimized network energy saving decisions, NG-RAN may need following information as input data for AI/ML-based network energy saving:

Input Information from Local node:

* UE mobility/trajectory prediction
* Current/Predicted Energy efficiency
* Current/Predicted resource status

Input Information from UE:

* UE measurement report (e.g. UE RSRP, RSRQ, SINR measurement, etc)

Input from neighbouring NG-RAN nodes:

* Past handover performance information
* Current/Predicted Energy efficiency
* Current/Predicted resource status
* Predicted Energy Saving Strategy (it is FFS which exact strategy will be exchanged)

If existing UE measurements are needed by a gNB for AI/ML-based network energy saving, RAN3 shall reuse the existing framework (including MDT and RRM measurements). FFS on whether new UE measurements are needed.

Editor’s Note: FFS other input information required for AI/ML-based network energy saving. FFS energy information is exact energy consumption value or energy efficiency gain.

#### 5.1.2.4 Output of AI/ML-based Network Energy Saving

AI/ML-based network energy saving model can generate following information as output:

* Energy saving strategies, e.g. recommended cell activation/deactivation, etc
* Predicted Energy Saving Strategy (it is FFS which exact strategy will be exchanged)
* Handover strategy, including recommended candidate cells for taking over the traffic
* Energy efficiency information, including predicted energy consumption information
* Predicted energy state (e.g. active, high, low, inactive)
* Validity time of energy strategy

Editor’s Note: FFS other output information expected from AI/ML-based network energy saving. FFS detailed granularity and action of energy saving strategy. FFS on accuracy of predicted energy saving decision.

#### 5.1.2.5 Feedback of AI/ML-based Network Energy Saving

To optimize the performance of AI/ML-based network energy saving model, following feedback can be considered to be collected from NG-RAN nodes:

* Load measurement
* Resource status of neighbouring NG-RAN node
* Energy efficiency
* UE transmission performance after handover

Editor’s Note: FFS other feedback expected from AI/ML-based network energy saving.

//////////////////////////////////////////////////////////////irrelevant operations skipped/////////////////////////////////////////////////////////////////////

Annex X (informative):

Overall structure of AI/ML based energy saving



Figure 1. Overall structure of AI-based energy saving

* Data collection is mainly responsible for the data collection, and required data is used to energy saving including current energy information, resource status of ES-cell and neighboring cells, even weather or other required information as input.
* AI-generated information includes load prediction, energy prediction, trajectory prediction, scenario classification or other required information generated by AI/ML model.
* Energy saving decision generation is to generate the energy saving decision based on the input from data collection, and refined input data. AI/ML model can also be used to make energy saving decision.
* Energy saving decision execution is to update the configuration based on the energy saving decisions and execute the decision.

 -----------------------------------End of Changes-----------------------------------