3GPP TSG-RAN WG3 Meeting #113-e R3-21xxxx

16-26 Aug 2021, E-meeting

**Title:** **[Draft]** LS to SA5 on model deployment and update from OAM to NG-RAN

**Response to:**

**Release:** Rel-17

**Work Item:** FS\_NR\_ENDC\_data\_collection

**Source:** **Intel [to be RAN3]**

**To:** SA5

**Cc:** SA2

**Contact person:**

* **Name: Ziyi Li**
* **E-mail Address: ziyi.li@intel.com**

**Send any reply LS to: 3GPP Liaisons Coordinator,** [**mailto:3GPPLiaison@etsi.org**](mailto:3GPPLiaison@etsi.org)

# 1 Overall description

RAN3 is currently evaluating the following RAN intelligence functional framework to support AI/ML capability at NG-RAN, see Figure 1. It is possible that “Model training” and “Model inference” are deployed at different network entities. In this case, network nodes which has the functionality of “Model training” may deploy/update well-trained AI/ML model to the network nodes which performs “Model inference”. Hence, when “Model training” is deployed at OAM and “Model inference” is deployed at NG-RAN, it requires OAM to deploy and update AI/ML model for NG-RAN to perform model inference.



Figure 1. RAN Intelligent Functional Framework

Additionally, RAN3 has agreed to focus on three use cases in Rel-17, i.e. load balancing, energy saving and mobility optimization. From RAN3 understanding, SA5 also defines solutions of above use cases in MDA (Management Data Analytics) and provide analytics report to MDAS (Management Data Analytics Service) consumers (including NG-RAN). As in Rel-17 SI “FS\_NR\_ENDC\_data\_collection”, NG-RAN may also have the capability of AI/ML, it is possible that NG-RAN to take well-trained model from OAM and use it to generate inference result(s), which is also proposed in RAN3 solutions to load balancing, energy saving and mobility optimization.

RAN3 would like to ask SA5 to provide feedback on following questions:

Whether OAM can support AI/ML life cycle management and related configuration management when Model training is in OAM and Model inference in NG-RAN? For example:

1. Whether it is possible to download trained model from OAM to NG-RAN?
2. Whether OAM is possible to provide model update to NG-RAN?
3. If yes for question 1) and 2), does OAM require model performance feedback from NG-RAN? What information is expected to collect from NG-RAN as model performance feedback for load balancing, energy saving and mobility optimization?

# 2 Actions

**To: SA5**

**ACTION: RAN3 kindly asks SA5 to provide answers for the questions and additional details that RAN3 shall further consider.**

# 3 Dates of next TSG RAN3 meetings

RAN3#114e Nov. 1~12, 2021 online

RAN3#115e Feb. 21~25, 2022 TBD