3GPP TSG RAN#102

Edinburgh , GB, December 11th -15th, 2023

RP-233554

Views on Rel-19 NG-RAN AI/ML

Agenda Item: 9.1.3.1

Source: Xiaomi

Document for: Discussion

Backgrounds

RAN#101 summary (RP-232623)

- **Rel-18 leftovers (depends on Rel-18 outcome)**
  - Mobility optimization
    - NR-DC
    - UE Trajectory prediction
  - Energy saving
    - Prediction for energy cost
  - Continuous MDT
- **At least 1 potential new use case/scenario which benefits from inference-based techniques**
  - Coverage and Capacity Optimization
  - Potential Rel-18 use case enhancements (e.g. new scenarios for ES)
  - QoE
  - Slicing
  - **Note:** Model training/inference at the DU and/or F1 enhancements must be justified by the use case
- **Study:** Consider available functionality specified in 5GC (NWDAF, MDAS) in order to align AI/ML framework between RAN and CN

RAN chair suggestion(RP-232745)

6 months SI + 12 months WI

At most 2 potential new use case/scenario which benefits from inference-based techniques

Rel-18 leftovers

Mobility optimization

- **NR-DC:** NR-DC scenario had been studied in R17, and NR-DC mobility is the main use case for mobility optimization.
- **UE Trajectory prediction:** in R18, only cell-level UE trajectory information is introduced, we believe a finer granularity UE trajectory information can be very helpful not only for UE mobility enhancement but also for other use cases e.g. load balance, CCO, etc.
- **Note: the overlaps between RAN3-led AI NG-RAN and RAN2-led AI mobility should be avoided.**

Continuous MDT

MDT can be a main tool for AI related data collection no matter for AI NG-RAN or AI Air, continuous MDT needs to be supported so that continuous data can be provided to the AI related functions.

Coordination with core network

- The NWDAF can provide finer granularity UE trajectory information, which can be used by RAN.
- the solution to support Continuous MDT may have SA2/SA5 involvement.

Proposal 1, the following R18 leftovers can be further considered in R19: Mobility optimization for NR-DC, UE Trajectory prediction, Continuous MDT, coordination with SA2/SA5 may be needed. The overlaps between RAN3-led AI NG-RAN and RAN2-led AI mobility should be avoided.

New use case (Coverage and Capacity Optimization)

Motivation:

- CCO is the main use case in network optimization.
- Good coverage and capacity design can enhance UE experience and reduce user complaints.
- Existing CCO have many challenges, it can only solve the issue when it occurs, considering the dynamic traffic and UE behaviors, it's hard to prevent the coverage or capacity issues in advance.
- To address these challenges, machine learning algorithms and predictive analytics can be helpful to continuously monitor and adjust the network for optimal performance.

Coordination with core network

- NWDAF can provide UE trends information including traffic and trajectory, which may be helpful for CCO.

Proposal 2, the AI-based Coverage and Capacity Optimization can be studied in R19, coordination with SA2 may be needed.

New use case (QoE)

Motivation:

- QoE indicates the overall level of customer satisfaction with a service, particularly in the digital realm such as internet services, mobile networks, and various types of applications.
- It had been discussed and agreed that RAN visible QoE can be helpful and useful for resource allocation in RAN.
- However, not all the UEs support RVQoE reporting or can report RVQoE.
- Existing RVQoE metrics are limited and the reporting mechanism is not efficient.
- Using AI to predict the QoE can bring benefits to many cases e.g. mobility optimization, load balance and resource allocation to ensure the UE experience.

Coordination with core network

- NWDAF can already provide predicted QoE of the UE from AF perspective, which can be used for RAN, and NWDAF can also provide well trained QoE model to RAN for real time optimization.
- SA5 may be involved if there’re enhancements on existing QoE collection mechanism for AI data collection purpose.

Proposal 3, the AI-based QoE can be studied in R19, coordination with SA2/SA5 may be needed.

Impacts on RAN split architecture

Consideration:

- The placement of the AI model inference can be discussed case by case. E.g. if the purpose is for dynamic scheduling, it's more efficient to deploy the AI model inference in gNB-DU.
- Currently, RAN1-led AI for Air Interface includes the uses cases of CSI, beam management and positioning, if RAN split architecture is considered, the gNB-DU is also more appropriate to perform the model inference.
- If RAN split architecture is supported in both RAN3-led AI/ML and RAN1-led AI/ML, whether a general mechanism or case by case mechanism is considered is a question.

Proposal 4, if RAN split architecture is supported in both RAN3-led AI/ML and RAN1-led AI/ML, it is suggested to discuss whether a general mechanism is considered.

Summary of R19 NG-RAN AI/ML Proposals

The following proposals are provided in with regard for R19 NG-RAN AI/ML:

Proposal 1, the following R18 leftovers can be further considered in R19: Mobility optimization for NR-DC, UE Trajectory prediction, Continuous MDT, coordination with SA2/SA5 may be needed. The overlaps between RAN3-led AI NG-RAN and RAN2-led AI mobility should be avoided.

Proposal 2, the AI-based Coverage and Capacity Optimization can be studied in R19, coordination with SA2 may be needed.

Proposal 3, the AI-based QoE can be studied in R19, coordination with SA2/SA5 may be needed.

Proposal 4, if RAN split architecture is supported in both RAN3-led AI/ML and RAN1-led AI/ML, it is suggested to discuss whether a general mechanism is considered.

Thanks