TSG RAN#100
Taipei, June 12 – 14, 2023

RP-231089

Data Collection for Offline Model Training at UE-side or Neutral site

Source: Qualcomm Incorporated, MediaTek Inc.

© 3GPP 2023

Summary of RAN1&RAN2 discussions on UE Model Training

- Based on RAN1#112 agreement, the model training can be performed at the UE-side or neutral site.
- Motivation
  - The training cannot practically happen in the UE for the lack of training environment including but not limited to availability of data, data storage capacity, computational capacity, and compilation capabilities.
  - RAN2 has agreed that only offline training is supported in Rel-18.
- SA2 involvement is important to ensure a coherent system-wide design (while taking into account RAN WGs findings)

Agreement

To facilitate the discussion, consider at least the following Cases for model delivery/transfer to UE, training location, and model delivery/transfer format combinations for UE-side models and UE-part of two-sided models.

Case	Model delivery/transfer	Model storage location	Training location
y	model delivery (if needed) over-the-top	Outside 3gpp Network	UE-side / NW-side / neutral site
z1	model transfer in proprietary format	3GPP Network	UE-side / neutral site
z2	model transfer in proprietary format	3GPP Network	NW-side
z3	model transfer in open format	3GPP Network	UE-side / neutral site
z4	model transfer in open format of a known model structure at UE	3GPP Network	NW-side
z5	model transfer in open format of an unknown model structure at UE	3GPP Network	NW-side

Note: The Case definition is only for the purpose of facilitating discussion and does not imply applicability, feasibility, entity mapping, architecture, signalling nor any prioritization.
Note: The Case definition is NOT intended to introduce sub-levels of Level z.
Note: Other cases may be included further upon interest from companies.
FFS: Z4 and Z5 boundary

RAN2#121bis-e agreement:

"R2 will deprioritize aspects of on-line/real-time training for the whole SI (unless R1 identifies that it is needed for one of the studied use cases)."

Definitions

- “UE-side training location” is a training entity under the control of the UE vendor.
  - It is NOT training at the UE device.

- “Neutral site training location” is a neutral training entity that is not under the control of the UE vendor or network vendors.
  - The “neutral entity” can be an operator or any third-party authorized and monitored by the operator, able to train models for a particular UE.

Requirements

- For offline model training at the UE-side and neutral site, the data collection solution should:
  - Allow data to reach “the UE-side or a neutral site” (ie Over The Top)
  - Collect a large amount of data, sufficient for model generalization
  - Collect non-standardized data for various models tailored to UE internal/external conditions and design choices (e.g. UE resource constraints, radio environment, feature engineering)

- An SA2 defined framework of “UE data collection” is more flexible while maintaining operator’s capability to authorize and monitor the data collection, and therefore better suited for that purpose

Data collection requirements for offline model training at the UE-side or at a Neutral site

- Proposal 1: a standard data collection solution is required to enable offline model training at the UE side, or at a Neutral site (with support for both).
- Proposal 2: The standard data collection solution shall:
  - Terminate at a network server allowing data to reach “the UE-side or a neutral site”
  - Be capable of collecting a large amount of data sufficient for model generalization
  - Be capable of collecting non-standardized data for various models tailored to UE internal/external conditions and design choices
- Proposal 3: RAN plenary to liaise SA/SA2 to recommend studying a standard data collection solution that enables an offline model training at the UE-side or a neutral-site in Rel-19.
  - Draft LS to SA,SA2 is submitted in RP-231091
  - Inputs from RAN study should be taken into account