3GPP TSG-RAN WG2 Meeting #126 R2-24xxxxx

Fukuoka, Japan, May 20th -24th, 2024

Agenda Item: 8.1.2.2

Source: Intel, Samsung

Title: Report of [POST126][032][AI/ML PHY] LCM (Intel/Samsung)

Document for: Discussion, Decision

# Introduction

This report provides a summary for the following post-meeting email discussion:

* [POST126][032][AI/ML PHY] LCM (Intel/Samsung)

Intended outcome:

Phase 1: Agreable definitions (Samsung)

Phase 2: Reach common understanding of reactive/proactive framework for applicable functionality.

Deadline: long

The deadline for providing comments for phase 1 is June 11, 2024 Tuesday at 21:00 UTC.

Companies providing input to this email discussion are requested to leave contact information below.

|  |  |  |
| --- | --- | --- |
| **Company** | **Name** | **Email Address** |
| Apple | Peng Cheng | Pcheng24@apple.com |

# Phase 1: definition of functionalities

In RAN2 #126 meeting, RAN2 discussed the following definition for functionality types and decided to have more discussion to identify the need of such definitions and whether further update is needed to clarify the definition [1].

|  |
| --- |
| Proposal 2: RAN2 agree the following definition for functionality types as a starting point.  - *Supported/identified functionalities:* this refers to functionalities that UE can indicate by using UE capabilities.  - *Configured functionalities:* this refers to functionalities that gNB can configure UE for model inference and performing measurements for training purposes?. Depending on proactive/reactive approach, configured functionalities may or may not be applicable upon configuration.  - *Applicable functionalities:* this refers to functionalities that the UE is ready to apply for model inference. It can be considered as candidates for functionality activation.  - *Activated functionalities:* this refers to functionalities that the UE starts predicting beam results via model inference. |

In this discussion, it would be good to discuss each functionality type.

## Supported functionalities

|  |
| --- |
| *Supported functionalities:* this refers to functionalities that UE can indicate by using UE capabilities. |

The moderator think that we can simplify the name from supported/identified functionalities to supported functionalities for convenience of discussion. Please comment if it is not ok.

**Q1: Do you agree that supported functionalities refer to functionalities that UE can indicate by using UE capability signaling and gNB/LMF can configure?**

If it is NO (or partially Yes), please add preferred definition (or additional definition) for it.

|  |  |  |
| --- | --- | --- |
| Company | Yes/No | Comment |
| Apple | Partial Yes | We prefer not to couple the “supported functionalities” with “configured functionalities” (i.e. the 2nd half sentence “and gNB/LMF can configure” can be removed):   1. As discussed in Q2/Q3, it is not clear whether the “gNB/LMF can configure” means “configure for training” and/or “configure for inference”. This may bring additional ambiguity. 2. Since UE capability is RAN2 expertise, we think RAN2 have common understand what “indicate by using UE capability signalling” means (i.e. UE supports the feature and NW can configure the feature). So, the 2nd half sentence is redundant.   Thus, we suggest below change:  **supported functionalities refer to functionalities that UE can indicate by using UE capability signaling ~~and gNB/LMF can configure~~** |
|  |  |  |

## Configured functionalities

|  |
| --- |
| *Configured functionalities:* this refers to functionalities that gNB/LMF configured to UE. UE for model inference and performing measurements for training purposes? Depending on proactive/reactive approach, configured functionalities may or may not be applicable upon configuration. |

As commented during RAN2 discussion, the moderator changed “gNB can configure” to “gNB configured”. And, LMF is added to cover positioning use case. Please comment if the change is not acceptable.

Please note that the second part (“Depending on proactive/reactive approach, configured functionalities may or may not be applicable upon configuration.”) will be discussed in Q5 in Section 2.3.

**Q2: Do you agree that configured functionalities refer to functionalities that gNB/LMF configured to UE?**

If it is NO (or partially Yes), please add preferred definition (or additional definition) for it.

|  |  |  |
| --- | --- | --- |
| Company | Yes/No | Comment |
| Apple | No | We doubt whether RAN2 really need this definition:   1. In 3GPP, we only specify UE behaviour under NW configuration. Thus, the definition suggested by moderator is obvious and doesn’t bring useful information. 2. As Q3 discussed, the configured functionalities may be understood as inference configuration or training configuration (and even performance monitoring configuration, although it was not mentioned in online discussion). So, the definition suggested by moderator may bring further ambiguity or further clarification. 3. We think at least stage 3 specification (either RRC or MAC) doesn’t need this definition because what NW configured to UE is always clearly specified in RRC and MAC spec. Thus, whether it is configuration for training or inference or monitoring should be crystal clear from text before and after in stage 3 specification. 4. In our understanding, the key controversial issue (or AI/ML specific issue) is the boundary among the following 3 terms: “supported functionality”, “applicable functionality” and “activated functionality”. Thus, it seems sufficient to define these 3 terms.   Thus, we think RAN2 don’t need definition of “configured functionalities”. |
|  |  |  |

**Q3: do you agree that this can be used for both UE-side model inference and training purpose?**

If it is NO, please indicate your preference e.g. only model inference for now or any suggestion for RAN2 progress.

|  |  |  |
| --- | --- | --- |
| Company | Yes/No | Comment |
| Apple |  | See our comment to Q2. |
|  |  |  |

## Applicable functionalities

|  |
| --- |
| *Applicable functionalities:* this refers to functionalities that the UE is ready to apply for model inference. It can be considered as candidates for functionality activation. |

It is also noted that RAN2 made a following agreement regarding applicable functionalities.

|  |
| --- |
| The UE will indicate the gNB/LMF whether the AI/ML functionality is available/applicable. For a functionality to be applicable at least there should at least one model available within it. FFS other details on what is applicability/non-applicability. |

**Q4: Do you agree that applicable functionalities refer to functionalities that the UE is ready to apply for model inference and they can be considered as candidates for functionality activation/deactivation?**

If it is NO (or partially Yes), please add preferred definition (or additional definition) for it.

|  |  |  |
| --- | --- | --- |
| Company | Yes/No | Comment |
| Apple | Yes with comments | We understand that “the UE is ready to apply for model inference” means the UE:   1. It has successfully completed model training compliant with UE’s internal status. And the trained model is already in UE device. 2. The dataset which was used for training the model is consistent with current inference configuration.    * Note that if it is not consistent, the UE can’t be regarded as “ready to apply” because current gNB is not aligned with UE and so can’t activate it.     Although we understand these are details which will be discussed in phase 2, it is better to confirm current definition can cover them. |
|  |  |  |

In relation to configured functionalities, configured functionalities may or may not be applicable immediately upon configuration depending on proactive/reactive approach. For example, in one way, UE already provides applicable functionalities/applicability related information and gNB can configure applicable functionalities. The other way would be that UE provide applicable functionalities/applicability related information after receiving configured functionalities from gNB and hence, configured functionalities may not be applicable immediately upon configured functionalities. We can remove “depending on proactive/reactive approach” in the definition as it will be discussed further in Phase 2 and there is no need to add as a part of definition.

**Q5: do you agree that configured functionalities may or may not be applicable immediately upon configuration?**

If it is NO (or partially Yes), please comment what relationship you envision between configured functionalities and applicable functionalities.

|  |  |  |
| --- | --- | --- |
| Company | Yes/No | Comment |
| Apple |  | See our comment to Q2. We think RAN2 only need to define “**supported functionalities**”, “**applicable functionalities**” and “**activated functionalities**”, i.e. no need to define “**configured functionalities**”. |
|  |  |  |

## Activated functionalities

|  |
| --- |
| *Activated functionalities:* this refers to functionalities that the UE starts predicting beam results via model inference. |

**Q6: Do you agree that activated functionalities refer to the functionalities that the UE uses beam prediction/positioning via model inference?**

If it is NO (or partially Yes), please add preferred definition (or additional definition) for it.

|  |  |  |
| --- | --- | --- |
| Company | Yes/No | Comment |
| Apple | partially Yes | We think “**UE uses beam prediction/positioning”** is not necessary, especially considering CSI compression/prediction is still on the table of Rel-19. Since it is a high level definition, we prefer to make it more general, e.g.  **activated functionalities refer to the functionalities that the UE uses ~~beam prediction/positioning via~~ output of model inference** |
|  |  |  |

## Available functionalities

During RAN2 discussion, there is a proposal on availability/available functionalities [2].

|  |
| --- |
| *Proposal 2 Introduce signalling for the UE to inform the gNB whether the AI/ML functionality is available for operation (e.g., whether there are trained models available within it). FFS whether the “availability indication” can be reported as part of the applicability-reporting information, or as a separate signalling.* |

**Q7: Do you agree that available functionalities should be separately considered from applicable functionalities?**

If yes, please comment what separate characteristics should be added/expected compared to applicable functionalities.

If no, please suggest how to consider available functionalities in RAN2 discussion (e.g. consider same as applicable functionalities or postpone to the next meeting)

|  |  |  |
| --- | --- | --- |
| Company | Yes/No | Comment |
| Apple | No (same as applicable functionalities) | Please note that above proposal was agreed as different wording after online discussion:  3 The UE will indicate the gNB/LMF whether the AI/ML functionality is available/applicable. For a functionality to be applicable at least there should at least one model available within it. FFS other details on what is applicability/non-applicability.  So, our understanding is that above agreement (esp. the highlighted part) has implied “applicable functionality” is same as “available functionality”.  Meanwhile, we also understand that the highlighted part implies the definition of “applicable functionalities suggested by moderator:  **applicable functionalities refer to functionalities that the UE is ready to apply for model inference and they can be considered as candidates for functionality activation/deactivation** |
|  |  |  |

# Phase 2

# Conclusion

# Reference

[1] R2-2405180 Functionality-based LCM for UE sided model Samsung

[2] R2-2405266 LCM for UE-side models for beam management Ericsson