3GPP TSG RAN WG1 #109-e R1-220xxxx

e-Meeting, May 9th – 20th, 2022

Source: Moderator (OPPO)

Title: Discussion summary#1 for other aspects on AI/ML for beam management

Agenda Item: 9.2.3.2

Document for: Discussion and Decision

# Introduction

The Rel-18 WID of AI/ML for NR Air Interface focuses on a subset of three typical use cases:

1. CSI feedback enhancement
2. Beam management
3. Positioning accuracy improvement.

This document focuses on the other aspects of AI/ML for beam managements, including representative sub use cases and potential specification impact. The company proposals are summarized, and offline proposals drafted passed on company contributions.

Regarding the file names, companies are encouraged to follow the guidance of R1-2203012 (Page 16) as below:

|  |
| --- |
| * + - To avoid ending-up with too long file names and downloading/opening issues, the following naming convention is recommended:       * Keep the previous company’s name (only the most recent one) in the filename, e.g.         + 5/Summary-1-v000-Moderator (HW)         + 5/Summary-1-v001-LG         + 5/Summary-1-v002-LG-CATT         + 5/Summary-1-v003-CATT-vivo         + 5/Summary-1-v004-Moderator(HW)       * It helps identifying on which previous version your input is based on and solve any crossing emails issue. Note the use of 3digit version numbers in the file names. |

# Contact Information

Respondents to the email discussion are encouraged to fill in the following table.

|  |  |  |
| --- | --- | --- |
| Company | Name | Email |
| Moderator | Zhihua SHI | szh@oppo.com |
| Apple | Yushu Zhang | yushu\_zhang@apple.com |
| vivo | Peng SUN | sunpeng@vivo.com |
| AT&T | Thomas Novlan | thomas\_novlan@labs.att.com |
| Futurewei | Baoling Sheen | bsheen@futurewei.com |
|  |  |  |
|  |  |  |

# Summary of Contributions and Offline Proposals

## Sub use cases

The objective on the use case of beam management is captured in R18 SID (RP-213599) as below:

|  |
| --- |
| Use cases to focus on:   * Initial set of use cases includes:   + CSI feedback enhancement, e.g., overhead reduction, improved accuracy, prediction [RAN1]   + **Beam management, e.g., beam prediction in time, and/or spatial domain for overhead and latency reduction, beam selection accuracy improvement [RAN1]**   + Positioning accuracy enhancements for different scenarios including, e.g., those with heavy NLOS conditions [RAN1] * Finalize representative sub use cases for each use case for characterization and baseline performance evaluations by RAN#98   + The AI/ML approaches for the selected sub use cases need to be diverse enough to support various requirements on the gNB-UE collaboration levels |

Following the SID, companies proposed a dozen of sub use cases for AI/ML-based beam management with different inputs, different outputs, different functionalities, different benefits and so on. The detailed observations and proposals are collected in Section 4.

### Categories and typical sub use cases

In order to facilitate the subsequent discussions, we categorize the diverse sub use cases, proposed by all the contributions of RAN1#109e, into the following types:

* Cat1: Spatial-domain DL beam prediction
  + **BM-Case1:** Spatial-domain DL beam prediction for Set A of beams based on measurement results of Set B of beams
* Cat2: Time-domain DL beam prediction
  + **BM-Case2:** Temporal DL beam prediction for Set A of beams based on the historic measurement results of Set B of beams
* Cat3: Others
  + **BM-Case3:** Beam prediction for higher frequency band (e.g., a band in FR2) based on measurement results of lower frequency band(s) (e.g., a band in FR1)
  + **BM-Case4:** Beam prediction based on UE positioning/trajectory
  + **BM-Case5:** Beam prediction in terms of Qos
  + **BM-Case6:** Spatial-domain UL beam prediction for Set A of beams based on measurement results of Set B of beams
  + **BM-Case7:** beam measurement feedback compression
  + **BM-Case8:** Parameter optimization to improve performance of multi-beam system

For **BM-Case1**, the main idea is that AI/ML model is to predict the top-N1 DL beams out of Set A based on the measurement results of Set B of beams. There were different options proposed by contributions for Set A and Set B and companies’ views are as below:

* Set B is a sub set of Set A.
  + Huawei [1], ZTE [2], Ericsson [3], IDC[4], CATT [5], vivo [6], NEC [7], Xiaomi [9], Samsung[10], OPPO[11], Beijing Jiaotong University[12], Panasonic [13], FUTUREWEI[14], CIACT[16], Apple[17], CMCC[18], Lenovo[20], Spreadtrum[21], TCL[22], Nokia[23], Intel[24], NVIDIA[25], Mavenir [27], QC[28], Fujitsu[29]
* Set A consists of narrow beams whereas Set B consists of wide beams
  + CATT [5], vivo [6], DOCOMO[19], Nokia[23], QC[28]

When N1 > 1, a second stage may be needed, i.e., the best beam is further selected based on the sweeping of AI/ML-predicted N1 beams. While most companies discussed the DL Tx beam predictions, there was also some companies discussing the DL Rx beam predictions, e.g., SS[10], Intel[24].

For **BM-Case2**, the main idea is that AL/ML model is to predict the information of DL beam(s) for future time based on historic measurement results. There were different proposals based on the contributions for the information of beam(s):

* Top-N2 beams and the predicted L1-RSRP
  + Huawei [1], ZTE [2], Ericsson [3], IDC[4], CATT [5], vivo [6], NEC [7], Sony[8], Xiaomi [9] (lower priority compared to Case 1), Samsung[10], OPPO[11] (lower priority compared to Case 1), Panasonic [13], FUTUREWEI[14], LGE[15] (high priority), Apple[17], DOCOMO[19], Spreadtrum[21], TCL[22], Nokia[23], Intel[24], NVIDIA[25], Mavenir [27], QC[28]
* Beam dwelling time
  + ZTE[2], NEC [7], Apple[17]
* Beam failure / blockage
  + Panasonic[13], TCL[22], QC[28]
* New candidate beam
  + Panasonic[13], TCL[22]

For **BM-Case3**, the main idea is that AL/ML model is to predict the top-N3 DL beams for a frequency band in FR2 based on the measurement results of a lower frequency band, which can be in FR1. A second stage is usually used to further the best beam by utilizing the sweeping of AI/ML-predicted N3 beams.

For **BM-Case4**, the main idea is that AL/ML model is to predict the best beams based on UE location/trajectory. For this case, new type of UE measurement/reporting, rather than beam management measurement/reporting, will be used.

For **BM-Case5**, the best beam is predicted according to the QoS class and its requirements. The main motivation is that UEs are sometimes scheduled to a beam that is not the strongest (signal strength wise) but a beam that can improve the UEs QoS metrics. Nokia [23] suggested reinforcement learning for this case.

For **BM-Case6**, the main idea is that AI/ML model is to predict the top-N6 UL beams out of Set A based on the measurement results of Set B of beams, where Set B is a subset of Set A. Generally speaking, BM-Case6 is the counterpart of BM-Case1 for the UL beam management.

For **BM-Case7**, the main idea is that UE generates the reported results by using AI/ML-based encoder (for compression) based on the beam measurement results and gNB recovers the beam information by using the corresponding AI/ML-based decoder.

For **BM-Case8**, one example in Mavenir[27] is that AI/ML model is to optimize the beam-specific cell individual offset (CFO) and Time-to-Trigger (TTT) parameters in real time or, near-real time. This would reduce the cases of too early, too late switches, ping-pong effects and switching failures and thus, reduce the latency and the outage rate. In general, it is to improve the beam-based mobility. Another example in Charter[30] is to optimize vector-quantized codebook for beam management on the gNB side.

Companies’ views are summarized in the following table:

Table 1: Sub use cases and categories

|  |  |  |
| --- | --- | --- |
| Category | Sub use case | Supported or mentioned (but doesn’t explicitly say no or low priority) by companies |
| Cat1:  Spatial-domain DL beam prediction | **BM-Case1:** Spatial-domain DL beam prediction for Set A of beams based on measurement results of Set B of beams | 26  Huawei [1], ZTE [2], Ericsson [3], IDC[4], CATT [5], vivo [6], NEC [7], Xiaomi [9], Samsung[10], OPPO[11], Beijing Jiaotong University[12], Panasonic [13], FUTUREWEI[14], CIACT[16], Apple[17], CMCC[18], DOCOMO[19], Lenovo[20], Spreadtrum[21], TCL[22], Nokia[23], Intel[24], NVIDIA[25], Mavenir [27], QC[28], Fujitsu[29] |
| Cat2:  Time-domain DL beam prediction | **BM-Case2:** Temporal DL beam prediction for Set A of beams based on the historic measurement results of Set B of beams | 22  Huawei [1], ZTE [2], Ericsson [3], IDC[4], CATT [5], vivo [6], NEC [7], Sony[8], Samsung[10], OPPO[11], Panasonic [13], FUTUREWEI[14], LGE[15], Apple[17], DOCOMO[19], Spreadtrum[21], TCL[22], Nokia[23], Intel[24], NVIDIA[25], Mavenir [27], QC[28], |
| Cat3: Others | **BM-Case3:** Beam prediction for higher frequency band (e.g., a band in FR2) based on measurement results of lower frequency band(s) (e.g., a band in FR1) | 2  Sony[8], Apple[17], |
| **BM-Case4:** Beam prediction based on UE positioning/trajectory | 2  Sony [8], Lenovo[20], |
| **BM-Case5:** Beam prediction in terms of Qos | 1  Nokia[23] |
| **BM-Case6:** Spatial-domain UL beam prediction for Set A of beams based on measurement results of Set B of beams | 1  Samsung[10], |
| **BM-Case7:** beam measurement feedback compression | 1  Samsung[10], |
| **BM-Case8:** The beam-specific parameter optimization | 2  Mavenir[27], Charter[30] |

Please provide your input with regard to the following aspects:

* Is any sub use case proposed in some tdoc(s) missing? If so, please add the related information including the brief description of the new sub use cases, the corresponding tdoc, and so on
* Is the position of some company misunderstood or wrongly captured? If so, please correct me.
* Descriptions of the above sub use cases
* …

|  |  |
| --- | --- |
| Company | Comments |
| Apple | We are fine to divide all use cases into two categories: 1) spatial domain beam prediction, 2) time domain beam prediction, 3) other. But it seems case 1, 3, 4, 5 are for spatial domain beam prediction, which should be under cat 1.  In addition, if case 2 includes beam dwelling time prediction, we suggest we explicitly mention that. From current formulation, it seems this only includes beam index prediction. |
|  |  |
|  |  |
|  |  |
|  |  |

As shown in the above table, majority companies support both **BM-Case1** and **BM-Case2**. Meanwhile, the other sub use cases are supported by a limited number of companies. Thus, it is suggested to focus on the sub use case **BM-Case1** and **BM-Case2** in the first step, and continue to discuss other sub use cases. The following proposal can be discussed, and further refined based on the inputs:

***Proposal 1-1: For AI/ML-based beam management, support BM-Case1 and BM-Case2 as the representative sub use case for characterization and baseline performance evaluations***

* ***BM-Case1: Spatial-domain DL beam prediction for Set A of beams based on measurement results of Set B of beams***
* ***BM-Case2: Temporal DL beam prediction for Set A of beams based on the historic measurement results of Set B of beams***
* ***FFS: details of BM-Case1 and BM-Case2***
* ***FFS: other sub use cases***

Please provide your input wrt description of the proposal as well as any other potential sub use cases that should be treated with high priority. In addition, please mention any other aspect that should be considered/included.

|  |  |
| --- | --- |
| Company | Comments |
| Apple | OK |
| vivo | Support |
| AT&T | support |
| Futurewei | We are ok to consider prioritizing BM-Case1 and BM-Case2 as 2 sub use cases for BM use case to facilitate the discussion. It is recommended to avoid too many sub use cases per each high level use case. At study/SI level, final set of representative sub use cases should be discussed and agreed across use cases. |
|  |  |

There were only a limited number of contributions to propose other sub use cases (i.e., BM-Case3, BM-Case4, BM-Case5, BM-Case6, BM-Case7, BM-Case8). However, since this is the first meeting, most of companies didn’t touch these sub use cases in their contributions. Thus, it would be good to check companies’ view to facilitate the decision how to proceed with these sub use cases.

***Collection of companies’ view:*** Companies are invited to input views on the following sub use cases and detailed explanation/ reasons are encouraged as well.

|  |  |  |
| --- | --- | --- |
| **Sub use cases for AI/ML based BM** | **Support** | **Not support** |
| BM-Case3 |  |  |
| BM-Case4 |  |  |
| BM-Case5 |  |  |
| BM-Case6 |  |  |
| BM-Case7 |  |  |
| BM-Case8 |  |  |

Detailed explanation/ reasons can be added to the following table

|  |  |
| --- | --- |
| Company | Comments |
| Apple | We support case 3. But since this is the first meeting, do we really need to exclude any use case? We suggest we list all cases for further study and companies can investigate the pros and cons. |
| AT&T | We are interested in Case 4 and Case 8 and agree that down selection is not needed at this phase. Characterization of the cases relative to Case 1 and Case 2 would especially be relevant (especially to the extent that common evaluation assumptions and AI/ML modeling approaches can be applied). |
| Futurewei | For the study in Rel-18, we suggest not considering BM-Case3 to BM-Case8 and focus on BM-Case1 and BM-Case2 only. |
|  |  |
|  |  |

### Details of sub use case **BM-Case1**

Where the AL/ML model is deployed is a key issue and has significant impact on the design and specification. Based on the contributions, there are different preferences for the AL/ML model deployment:

* AL/ML model deployed at NW side is preferred
* AL/ML model deployed at UE side is preferred
* both AL/ML model deployed at NW side and AL/ML model deployed at UE side are studied in R18
* Joint AL/ML model at NW and UE size can be studied

Additionally, AT&T[26] proposed to study centralized (e.g., across multiple gNBs) AI/ML operations. For the difference preferences, companies’ views are as below:

Table 2: AI model deployment

|  |  |  |
| --- | --- | --- |
|  | Preferred or mentioned | Not preferred |
| AI model deployed at NW side | Huawei [1], ZTE [2], Ericsson [3], IDC [4], CATT[5], Sony [8], Xiaomi[9], Samsung[10], LGE[15], CIACT[16], CMCC[18], DOCOMO[19], Spreadstrum[21], Nokia[23], AT&T[26], QC[28] |  |
| AI model deployed at UE side | Huawei [1], Ericsson [3], IDC [4], CATT[5], Sony [8], Xiaomi[9], Samsung[10], LGE[15], CAICT[16], CMCC[18], Spreadstrum[21], Nokia[23], AT&T[26], QC[28], Charter[30] | ZTE [2], |
| Joint AI at both NW and UE | Samsung[10], | Ericsson [3], |
| Joint AI across multiple gNB | AT&T[26] |  |
| Note:   * This table doesn’t differentiate the sub use cases. Some companies may prefer one deployment for one sub use case and prefer another deployment for other use cases. | | |

According to the above table, the first two types of AI/ML model deployment were supported or mentioned by most companies. Thus, the following proposal can be discussed, and further refined based on the inputs:

***Proposal 2-1: For the sub use case BM-Case1, down select one of the following AI/ML model deployments:***

* ***Alt.1: AI/ML models deployed at NW side***
* ***Alt.2: AI/ML models deployed at UE side***
* ***Alt.3: Both AI/ML models deployed at NW side and AI/ML models deployed at UE side***

Please provide your input wrt the above alternatives in the above proposal. In addition, feel free to provide other comment/suggestion

|  |  |
| --- | --- |
| Company | Comments |
| Apple | We are not sure whether we understand the definition of “deployed” correctly, especially for Alt3. Does it mean training in one side but inference is in the other side? |
| vivo | We prefer Alt 2, but we are open to study Alt 1 and Alt 3. |
| AT&T | We support Alt. 1 and are open to studying Alt. 2/3 as well |
| Futurewei | Alt.3. It is preferred to be flexible in deployment scenarios at this early stage of the SI. |
|  |  |

For an AI/ML-based mechanism, two basic factors are the input and output. Thus, in order to finalize a specific sub use case, the input and output should be clearly defined. As we discussed in Section 3.1.1, there are different proposals for the Set A and Set B. Thus, based on the contributions, the following proposal can be discussed, and further refined based on the inputs:

***Proposal 2-2: For the sub use case BM-Case1, down select one of the following alternatives:***

* ***Alt.1: Set B is a subset of Set A***
  + ***FFS: the number of beams in Set A and B***
  + ***FFS: how to determine Set B out of the beams in Set A (e.g., fixed pattern, random pattern, …)***
* ***Alt.2: Set A consists of narrow beams and Set B consists of wide beams***
  + ***FFS: the number of beams in Set A and B***
* ***Alt.3: both Alt1 and Alt.2***

Please provide your input wrt the above alternatives in the above proposal. In addition, feel free to provide other comment/suggestion. Further refinement will be based on the inputs.

|  |  |
| --- | --- |
| Company | Comments |
| Apple | We think Alt3 should be the most straightforward way. Both sub-sampling and hierarchical based approaches should be valid. We failed to see any reason to preclude one.  In addition, we suggest adding the following:   * ***Alt.4: set A and set B beams are in different bands*** |
| vivo | Alt 3 with the understanding that Alt 1 and Alt2 have the same core issue: finding an AI algorithm with good prediction performance gain and also provide sufficient flexibility on AI model deployment. |
| AT&T | We support Alt. 3 |
| Futurewei | Alt.3, be flexible and open to both options at the early stage of the SI. Down selection may be discussed/considered when evaluation results (including overhead) are available. |
|  |  |

For AI/ML input, there are also different proposals for companies. Thus, based on the contributions, the following proposal can be discussed, and further refined based on the inputs:

***Proposal 2-3: Regarding the sub use case BM-Case1, further study the following alternatives for AI/ML input with potential down-selection:***

* ***Alt.1: Only L1-RSRP measurement based on Set B of DL Tx beams***
* ***Alt.2: L1-RSRP measurement based on Set B of DL Tx beams and the corresponding beam ID***
* ***Alt.3: …***

Please provide your input wrt the above alternatives in the above proposal as well as any other potential alternative that should be treated with high priority. In addition, feel free to provide other comment/suggestion. Further refinement will be based on the inputs.

|  |  |
| --- | --- |
| Company | Comments |
| Apple | We suggest adding the following alternative:   * ***Alt.3: CIR based on Set B of DL Tx beam(s)*** |
| vivo | From our perspective, 3 alternatives can be studies,   * ***Alt 1: Only L1-RSRP measurement based on Set B of DL Tx beams*** * ***Alt 2: L1-RSRP measurement based on Set B of DL Tx beams and assistance information which may include, beam ID, beam angle or position information*** * ***Alt 3: L1-RSRP measurement based on Set B of DL Tx beams, assistance information and expected information which the output of AI model is predicted partial RSRPs corresponding to expected Rx angle in AI input.***   Thus, different assistance information can be discussed in Alt 2 or Alt 3. |
| Futurewei | Regarding the input, we prefer leaving it open to company’s implementation. |
|  |  |
|  |  |

***Proposal 2-4: Regarding the sub use case BM-Case1, further study the following alternatives for AI/ML output with potential down-selection:***

* ***Alt.1: Beam ID(s) and the predicted L1-RSRP of the predicted Top-N1 DL Tx beams*** 
  + ***FFS: N1***
* ***Alt.2: …***

Please provide your input wrt the above alternatives in the above proposal as well as any other potential alternative that should be treated with high priority. In addition, feel free to provide other comment/suggestion. Further refinement will be based on the inputs.

|  |  |
| --- | --- |
| Company | Comments |
| Apple | We suggest adding the following. Compared to predicted L1-RSRP in Alt1, the possibility for the beam to be the best beam could be more important.  ***Alt.2: Beam ID(s) of the predicted Top-N1 DL Tx beams***  ***Alt.3: Beam ID(s) and possibility for the beam to be the best beam of the predicted Top-N1 DL Tx beams*** |
| vivo | This is first meeting. We can make the categories more generic and inclusive.  ***Proposal 2-4: Regarding the sub use case BM-Case1, further study the following alternatives for AI/ML output with potential down-selection:***   * ***Alt.1: L1-RSRP*** * ***Alt.2: Beam information, such as beam ID, beam angle.*** * ***Alt 3: confidence level*** * ***Alt 4:…*** |
| Futurewei | We suggest leaving the output also open to company’s implementation choice as long as we agree on the set of performance evaluation metrics. For example, if companies agree to use Top-1/optimal/best beam prediction accuracy as one of the KPIs, then directly predicting L1-RSRP may not be required. If, however, companies agree to use measured RSRP or RSRP gap as one of the KPIs, then at least the predicted L1-RSRP (for the Top-1 or Top-K beams) has to be available from the output (either directly as a prediction output or not). |
|  |  |
|  |  |

There may be some other issues for each sub use cases. For example, whether online training or offline training is assumed, which is also related to the discussion/output of AI 9.2.1. We can discuss these issues later.

Please provide your input wrt any other issues that should be discussed with higher priority, any other suggestion/comment, …

|  |  |
| --- | --- |
| Company | Comments |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

### Details of sub use case **BM-Case2**

***Proposal 3-1: For the sub use case BM-Case2, down-select one of the following AI/ML model deployments:***

* ***Alt.1: AI/ML models deployed at NW side***
* ***Alt.2: AI/ML models deployed at UE side***
* ***Alt.3: Both AI/ML models deployed at NW side and AI/ML models deployed at UE side***

Please provide your input wrt the above alternatives in the above proposal. In addition, feel free to provide other comment/suggestion. Further refinement will be based on the inputs.

|  |  |
| --- | --- |
| Company | Comments |
| Apple | Similar to proposal 2-1, we are not sure whether we understand the definition of “deployed” correctly, especially for Alt3. Does it mean training in one side but inference is in the other side? |
| vivo | Similar to proposal 2-1 |
| AT&T | We support Alt. 1 and are open to studying Alt. 2/3 |
| Futurewei | Alt.3. It is preferred to be flexible in deployment scenarios at this early stage of the SI. |
|  |  |

There are different proposals from companies for the Set A and Set B. Thus, based on the contributions, the following proposal can be discussed, and further refined based on the inputs

***Proposal 3-2: For the sub use case BM-Case2, further study the following alternatives with potential down-selection:***

* ***Alt.1: Set B is a subset of Set A***
  + ***FFS: the number of beams in Set A and B***
  + ***FFS: how to determine Set B out of the beams in Set A (e.g., fixed pattern, random pattern, …)***
* ***Alt.2: Set A and Set B are the same***
  + ***FFS: the number of beams in Set A and B***
* ***Alt.3: …***

Please provide your input wrt the above alternatives in the above proposal as well as other potential alternatives should be with high priority. In addition, feel free to provide other comment/suggestion. Further refinement will be based on the inputs.

|  |  |
| --- | --- |
| Company | Comments |
| Apple | For beam prediction, it seems there is no need to define the connection between set A and set B. |
| vivo | Support. |
| AT&T | We support the proposal and prefer Alt. 2 |
| Futurewei | Alt.2 may be a more typical case in time domain beam prediction. Depending on evaluation results, Alt.1 may be ok as well. In summary, we prefer leaving this open at the early stage of SI. |
|  |  |

For the historic measurement results, it is natural to have a window or a number on the past measurement instances. Thus, the following proposal can be discussed, and further refined based on inputs.

***Proposal 3-3: Regarding the sub use case BM-Case2, the measurement results of K (K>=1) past measurement instances are used for AI/ML model input:***

* ***FFS: values of K***

Please provide your input wrt the description of the above proposal. In addition, feel free to provide other comment/suggestion. Further refinement will be based on the inputs.

|  |  |
| --- | --- |
| Company | Comments |
| Apple | OK in principle, but we think the K instances should be with the same interval, right? |
| vivo | Fine |
| AT&T | Ok |
| Futurewei | We prefer not fixing the K (K>=1)value and leave it to implementation decision. |
|  |  |

***Proposal 3-4: Regarding the sub use case BM-Case2, further study the following alternatives of measurement results for AI/ML input (for each past measurement instance) with potential down-selection:***

* ***Alt.1: L1-RSRP measurement based on Set B of DL Tx beams and the corresponding beam ID***
* ***Alt.2: …***

Please provide your input wrt the above alternatives in the above proposal as well as any other potential alternative that should be treated with high priority. In addition, feel free to provide other comment/suggestion. Further refinement will be based on the inputs.

|  |  |
| --- | --- |
| Company | Comments |
| Apple | We are a bit confused with the definition of set B, does it mean the reported beam(s) or the measured beam(s). If it includes reported beam(s), maybe beam index needs to be added as input in addition to L1-RSRP. |
| vivo | Similar to proposal 2-3. And we believe spatial domain beam prediction algorithm is an essential precondition for time domain beam prediction study. Thus, proposal 3-4 should be a supplement with time domain specific parameters on proposal 2-3. |
| Futurewei | Alt.1 is ok while companies are free to decide whether to use both as input to their AI/ML model. |
|  |  |
|  |  |

***Proposal 3-5: Regarding the sub use case BM-Case1, further study the following alternatives for AI/ML output (one prediction for a future time instance) with potential down-selection:***

* ***Alt.1: Beam ID(s) and the predicted L1-RSRP of the predicted Top-N2 DL Tx beams*** 
  + ***FFS: N1***
* ***Alt.2: Beam ID(s) and the corresponding beam dwelling time***
* ***Alt.3: Predicted Beam failure and the corresponding bream ID(s)***
* ***Alt.4. Predicted new candidate beam(s)***
* ***Alt.5. …***

Please provide your input wrt the above alternatives in the above proposal as well as any other potential alternative that should be treated with high priority. In addition, feel free to provide other comment/suggestion. Further refinement will be based on the inputs.

|  |  |
| --- | --- |
| Company | Comments |
| Apple | We suggest adding Alt 1b/1c as follows and change “Alt1” into “Alt1a”:   * ***Alt.1b: Beam ID(s) of the predicted Top-N2 DL Tx beams*** * ***Alt.1c: Beam ID(s) and possibility for the beam to be the best beam of the predicted Top-N2 DL Tx beams*** |
| vivo | Similar to proposal 2-4, and we suggest following modifications in proposal 3-5:  ***Proposal 3-5: Regarding the sub use case BM-Case2, further study the following alternatives for AI/ML output (one prediction for a future time instance) with potential down-selection:***   * ***Alt.1: L1-RSRP*** * ***Alt.2: Beam information, such as beam ID, beam angle.*** * ***Alt 3: confidence level*** * ***Alt 4: Beam dwelling time*** * ***Alt.5: Predicted Beam failure*** * ***Alt.6. Predicted new candidate beam(s)*** * ***Alt.7. …*** |
| AT&T | Ok the additional alternatives from vivo, although “confidence level” may need a more detailed definition (e.g. absolute, relative?) |
| Futurewei | Per description provided before, this question should be for BM-Case2.  Alt.1 for BM-Case2. However, we suggest modifying the wording a bit to **“*Beam ID(s) and/or the predicted L1-RSRP of the predicted Top-N2 DL Tx beams”***. We may discuss/decide whether both are needed once the EVM discussion is more stable/clearer. |
|  |  |

As the AI/ML model predicts the beam information for future time, it should be clear how many future time instances the prediction are made. Thus, the following proposal can be discussed, and further refined based on inputs.

***Proposal 3-6: Regarding the sub use case BM-Case2, AI/ML model output should be F predictions for F future time instances, where each prediction is for each time instance.***

* ***At least F = 1***
* ***FFS: other values of F***

Please provide your input wrt the description of the above proposal. In addition, feel free to provide other comment/suggestion. Further refinement will be based on the inputs.

|  |  |
| --- | --- |
| Company | Comments |
| Apple | OK in principle, but should the F instances with the same interval? |
| vivo | Fine |
| AT&T | Ok |
| Futurewei | At least 1 and let companies to decide how many future time instances should be. |
|  |  |

There may be some other issues for each sub use cases. For example, whether online training or offline training is assumed, which is also related to the discussion/output of AI 9.2.1. We can discuss these issues later.

Please provide your input wrt any other issues that should be discussed with higher priority, any other suggestion/comment, …

|  |  |
| --- | --- |
| Company | Comments |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

## Potential spec impacts

Generally speaking, the spec impacts heavily depend on the detailed sub use cases, e.g., some related aspects are as below:

* What type of training: online or offline?
* Where the AI/ML is deployed: at UE side, at NW side, at both UE and NW side?
* What the input is?
* What the output is?
* …

Thus, the spec impacts discussed in contributions are usually targeted to some specific sub use cases and the potential spec enhancements are quite diverse. Considering the group are still discussing what sub use cases should be supported, a brief summary is trying to capture the key aspects of potential spec enhancement, without many details:

* New or enhanced mechanism(s) to facilitate data collection for UE/NW model, e.g., training, fine-tuning, verification, e.g., some examples are mentioned by contributions
  + Enhanced BM procedures (including signalling/configuration, reporting) to facilitate the training data collection
  + Introduction of some new information, e.g., UE positioning, information from sensor (e.g., velocity, orientation, rotation)
  + Other assistance information for training
* New or enhanced mechanism(s) to facilitate AI/ML inference, e.g., some examples are mentioned by contributions
  + Enhanced BM measurement/reporting for AI inference
  + Signaling/configuration for enhanced BM measurement/reporting
  + Assistance information for AI inference
* New or enhanced mechanism(s) to facilitate AI model life cycle management, e.g., some examples are mentioned by contributions
  + Mechanisms/assistance information for AI/ML model activation, deactivation
  + Mechanisms/assistance information for AI model selection
  + Mechanisms/assistance information for Performance monitoring
  + May include the exchange of some assistance information
* AI-related UE capability and reporting
* Interface of AI model, e.g., input, output
* Other enhancements

This brief summary is not a complete list, and is just used for information. Please see Section 4 for more information. More details will be provided and more discussions will proceed according to the progress of sub use case discussion.

Please share your comment/suggestion on the discussion on spec impact.

|  |  |
| --- | --- |
| Company | Comments |
| Apple | We agree with FL that spec impact depends on use cases. This may be discussed after we identified the use cases. |
| Futurewei | We believe the discussions on specification impact can wait till the discussions on sub use cases and deployment options are more stable. |
|  |  |
|  |  |
|  |  |

# Detailed Proposals / Observations

|  |  |
| --- | --- |
| Huawei [1] | ***Observation 1: NW-oriented AI/ML beam management with the AI/ML model trained and inferred both at gNB side can operate under the collaboration level without AI/ML model exchange over the air-interface.***  ***Observation 2: UE-oriented AI/ML beam management with the AI/ML model trained and inferred both at UE side may require the NW to train the AI/ML model and deliver to the UE.***  ***Proposal 1****:* ***For AI/ML-based beam management, the following two sub use cases can be studied:***  ***Beam management in spatial domain***  ***Beam prediction in time domain***  ***Proposal 2: Study whether potential specification impact is needed for AI/ML-based beam prediction considering the following aspects:***  ***AI/ML model training procedure***  ***Enhancement for RSRP report and beam ID report***  ***AI/ML model monitoring procedure*** |
| ZTE [2] | ***Observation 1:*** *The traditional beam training method with brute-force sequential beam searching may result in excessive training overhead, measurement power consumption and processing delay.*  ***Observation 2:*** *Compared with the traditional exhaustive search-based beam training method, the learning-based beam prediction methods can significantly reduce the training overhead and processing delay.*  ***Proposal 1:*** *Both the AI/ML-based spatial-domain beam prediction and time-domain beam prediction should be taken into consideration for possible applicability.*  ***Proposal 2:*** *At least the scenario where AI model is deployed at base station should be considered for specification enhancements on beam management.*  ***Proposal 3:*** *Study potential enhancements on receive beam-related information to the base station reported by UE to assist gNB for more accurate beam prediction and improved generalization of the trained AI/ML model.*  ***Proposal 4:*** *Study potential enhancements to support more flexible beam measurement and reporting, in order to adapt to different AI/ML based beam prediction.*  ***Proposal 5:*** *Consider predictable mobility for beam management as an enhancement aspect for improving UE experience in FR2 high mobility scenario (e.g., high-speed train and high-way) in Rel-18 AI-PHY, which at least includes beam-management-related enhancements for predictable mobility, involving beam measurement, beam report and beam indication.* |
| Ericsson [3] | [Observation 1 Proprietary beam management procedures executed on the UE side (resp. NW side) effect NW side (resp. UE side) data quality and, therefore, AI/ML model generation and performance.](#_Toc102160598)  [Observation 2 Single-sided AI/ML models in the UE, which are transparent to the NW, may make beam management and AI/ML model generation more difficult.](#_Toc102160599)  [Observation 3 UE-side beam prediction AI/ML capability signalling can enable improving NW performance.](#_Toc102160600)  [Proposal 1 Study UE-sided AI enhancements for beam management.](#_Toc102160601)  [Proposal 2 Study NW-sided AI enhancements for beam management.](#_Toc102160602)  [Proposal 3 Do not study dual-sided joint AI enhancements for beam management.](#_Toc102160603)  [Proposal 4 Study enhanced beam management procedures to aid data collection for (offline) single-sided UE/NW model generation, for example potential assistance information.](#_Toc102160604)  [Proposal 5 Study specification impacts for beam prediction AI/ML model configuration activation, deactivation, and monitoring.](#_Toc102160605)  [Proposal 6 Study AI/ML model capability reporting associated with beam predictions from UE.](#_Toc102160606)  [Proposal 7 Study the benefit of signalling predicted values and associated confidence levels for beam management.](#_Toc102160607)  [Proposal 8 Quantify the benefits of site-specific beam prediction models AI/ML models trained on site-specific data.](#_Toc102160608)  [Proposal 9 Study requirements and solutions for enabling trained site-specific beam prediction AI/ML models to UEs](#_Toc102160609) |
| IDC [4] | ***Observation 1:*** *The current NR specification supporting UE reporting with up to 4 best CRIs/SSBRIs with L1-RSRP or L1-SINR can be very limited for gNB estimation.*  ***Observation 2:*** *Partial beam measurement which allows beam prediction by measuring only a subset of beams could be beneficial for reducing RS overheads and reporting latency.*  ***Observation 3:*** *For partial beam measurement, both UE side beam prediction and gNB side beam prediction have benefits.*  ***Observation 4:*** *The current NR specification supports measurement restriction to limit UE measurement, however, measurement restriction is to efficiently utilize RS transmissions for multiple beams not to consider time domain characteristics of beam measurement.*  ***Observation 5:*** *For gNB which predicts beams by using AI/ML, time domain characteristics of beam measurements are essential as well as spatial domain characteristics.*  ***Observation 6:*** *The current NR specification does not consider association between beams with different beam widths.*  ***Observation 7:*** *Utilizing association between beams with different beam widths can provide benefits for prediction accuracy e.g., robust estimation/identification of whole spatial characteristics with wide beams and accurate beam identification with narrow beams.*  ***Observation 8:*** *For Rel-15 beam management, actual mapping between DL Tx beam and UE Rx beam is totally based on UE implementation.*  ***Observation 9:*** *The implementation-based UE Rx beam selection works for Rel-15, however, UE Rx beam information is crucial to accurately predict beam qualities for AI/ML based beam prediction.*  ***Proposal 1:*** *Study benefits of simple specification extension of UE reporting such as* *increasing number of possible best CRIs/SSBRIs, introduction of absolute RSRPs for other CRIs/SSBRIs than the best and other possible extensions and other possible extensions*.  ***Proposal 2:*** *Study benefits of partial beam measurement for specification enhancement to reduce RS overheads and UE reporting latency*.  ***Proposal 3:*** *Study benefits of both UE side beam prediction and gNB side beam prediction for partial beam measurement*.  ***Proposal 4:*** *Study benefits of specification enhancements such as UE reporting with associated time slot domain information*.  ***Proposal 5:*** *Study benefits of specification enhancements on association between beams with different beam widths*.  ***Proposal 6:*** *Study benefits of specification enhancements on acquiring UE Rx beam information*. |
| CATT [5] | ***Proposal 1: The following sub use cases can be considered in Rel-18*** ***AI/ML-based beam management:***  ***Narrow beam prediction based on wide beam measurement.***  ***All beam prediction based on partial beam measurement.t***  ***Beam prediction in time domain.***  ***Proposal 2: The following spec impact of AI/ML based beam management can be considered:***  ***Signaling/procedure of AI model training/updating/fallback;***  ***Interface of AI model, i.e. relationship between measured RS and reported information;***  ***New procedure for RS measurement and reporting;***  ***Signaling/procedure design on exchanging AI-related/non-AI-related assistance information.*** |
| Vivo [6] | Two main sub use cases can be considered for spatial domain beam prediction to reduce overhead and/or improve beam searching accuracy.  Fine beam prediction based on coarse beam measurement  Super-resolution beam prediction based on partial beam measurement.  Study different variations for each sub-use case, considering generalization performance for different number of Tx and Rx beams.  The sub use case of time domain prediction is to predict future beam RSRP with historical RSRPs for both of the following purposes,  Time domain beam prediction for overhead reduction  Time domain beam prediction for accurate beam switching time  Spatial domain beam prediction algorithm is an essential precondition for time domain beam prediction study.  Study impact of different beam sweeping patterns for time domain beam prediction.  Study model deployment procedure and specification impact for both cases that beam prediction functionality resides in UE side and the functionality resides in gNB side.  Study sub-use cases from collaboration level 0~ level 3 for beam management cases.  Study configuration method of beam angle with minimum exposures of implementation details.  Study the gains and impact of different beam input orders.  Study the impact of input beam patterns to model performance monitoring for both spatial and time predictions. |
| NEC [7] | ***Observation 1: At least from the perspective of supporting various gNB-UE collaboration levels and having significant potential specification impact, beam prediction in spatial/time domain can be final representative sub use cases, while*** ***beam selection accuracy improvement can’t.***  ***Observation 2: In order to ensure the performance of the AI model in the real environment, verification of dataset including training, validation and testing is essential.***  ***Observation 3: For sub use case, multiple AI models may be arranged or deployed.***  ***Observation 4: For periodic or semi-persistent beam reporting, overhead of beam measurement and reporting resources can be reduced with beam prediction in time domain.***  ***Proposal 1: Support beam prediction in spatial/time domain as the final representative sub use cases.***  ***Proposal 2: Study the mechanism of exchanging information indicting verification results between gNB and UE.***  ***Proposal 3: Study the mechanism of reporting more beams, e.g., K>4.***  ***Proposal 4: Study the mechanism of model selection.***  ***Proposal 5: Study the mechanism of discontinuous reporting in periodic or semi-persistent beam reporting.***  ***Proposal 6: Study the method of indicating the future beam and the application time of beam.*** |
| Sony [8] | **: Support using AI/ML model trained with all locations and directions of UE for beam prediction and selection at gNB.**  **: Support using dynamic/aperiodic CSI-RS resource set to inform UE candidate beams.**  **: Beam prediction at gNB based on UE’s measurement report can be supported.**  **: Support beam prediction at gNB by using multi pieces of prior CSI information at low frequency for model training.**  **: Propagation environment based AI/ML model selections can be considered at gNB.** |
| Xiaomi[9] | ***Proposal 1: Study sub use case of beam prediction in spatial domain with high priority.***  ***Proposal 2: Prefer AI/ML model training at gNB side and the collaboration level Cat.2 since Rx beam information should be included in beam report for data set collection.***  ***Proposal 3: To discuss whether a common AI model or separate AI models will be trained for UE with different number of Rx beam.***  ***Proposal 4: To input beam information associated with each L1-RSRP to AI/ML model and study how to indicate the Tx beam information of gNB to UE for UE side inference.***  ***Proposal 5: To indicate Rx beam ID to UE for obtaining L1-RSRP input to AI/ML model.*** |
| Samsung[10] | **Proposal 1: Study the sub use cases for beam prediction in spatial domain.**  **Consider gNB-side beam prediction as a sub use case, and**  **Consider UE-side beam prediction as a sub use case.**  **Proposal 2: Study the sub use cases for beam prediction in time domain.**  **Consider gNB-side beam prediction as a sub use case, and**  **Consider UE-side beam prediction as a sub use case.**  **Proposal 3: Study the combination of spatial domain beam predication and time domain beam predication.**  **Proposal 4: Study beam measurement feedback compression as a candidate sub use case.** |
| OPPO[11] | ***Collaboration framework 0a and 0b involves no AI/ML-specific signaling nor model exchange, but they can enable AI/ML-based beam prediction with the aid of existing NR mechanism or modified/enhanced NR system;***  ***Collaboration framework 1b involves AI/ML-specific signaling but no model exchange and it facilitates inference at both sides for AI/ML beam prediction.***  ***For overhead and latency reduction, study the sub use case of AI/ML-based beam prediction in spatial domain as a starting point.***  ***Beam prediction in time domain can be studied and evaluated, but not with top priority.***  ***Study the mechanism and necessity of collaboration framework(s) on a per use case basis for AI/ML beam management.*** |
| Beijing Jiaotong University[12] | ***Proposal #1: Study*** ***image-reconstruction-based beam selection scheme as a use case for beam management enhancement.***  ***Proposal #2: Study the necessary specification change to support AI-based beam selection, considering collaboration level between UE and network, and additional signaling enhancement from the side of network for the adaptive beam measurement.*** |
| Panasonic[13] | **Observation 1: Initial beam establishment is one sub use case.**  **Observation 2: For beam tracking and refining, the following can be considered as sub use cases:**  **Adjustment of measurement/reporting interval**  **Predictive beam switching**  **Partial beam set measurement**  **Observation 3: For beam failure recovery, the AI/ML approaches would be similar to beam tracking and refining sub use cases.**  **Observation 4: For sub use case of initial beam establishment, all AI/ML functionalities located at UE can be considered as baseline, and it can be FFS to spread AI/ML functionalities between UE and network.**  **Observation 5: For sub use case of adjustment of measurement/reporting interval, network based AI/ML can be considered as baseline, and it can be FFS to spread AI/ML functionalities between UE and network.**  **Observation 6: For sub use case of predictive beam switching for RRC\_CONNECTED, network based AI/ML can be considered as baseline, and it can be FFS to spread AI/ML functionalities between UE and network.**  **Observation 7: For sub use case of partial beam measurement, both network based and UE based AI/ML can be considered as baseline, and it can be FFS to spread AI/ML functionalities between UE and network.**  The proposals are as follows:  **Proposal 1: AI/ML mapping within the network (such as gNB or OAM) is up to RAN2/3 discussion. RAN1 discussion should focuses network-UE relation.**  **Proposal 2: Consider the following mapping between sub use cases and network-UE collaboration levels for further study:**   |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | | **Sub use cases** | **Cat-1-UE**  (allAI/ML functionalities at UE) | **Cat-1-network**  (allAI/ML functionalities at network) | **Cat-2**  (Data Collection, Model Training and Model Inference at network; Actor at UE) | **Cat-3**  (Date Collection at network; Model Training, Model Inference and Actor at UE) | **Cat-4**  (Date collection and Model training at network; Model Inference and Actor at UE) | **Cat-5**  (Model Training and Model Inference at both network and UE ) | | **Initial beam establishment** | **Baseline** | **Deprioritzed** | **Deprioritized** | **FFS** | **FFS** | **FFS** | | **Adjustment of measurement/reporting interval** | **FFS** | **Baseline** | **Baseline** | **FFS** | **FFS** | **FFS** | | **Predictive beam switching** | **FFS** | **Baseline** | **Baseline** | **FFS** | **FFS** | **FFS** | | **Partial beam set measurement** | **Baseline** | **Baseline** | **Baseline** | **Baseline** | **FFS** | **FFS** | |
| FUTUREWEI[14] | ***Observation 1: AI/ML-based beam prediction in spatial domain and AI/ML-based beam prediction in time domain can potentially reduce the overhead, improve beam selection accuracy, and improve UE experience/performance.***  ***Proposal 1: Support “AI/ML-based beam prediction in spatial domain” and “AI/ML-based beam prediction in time domain” as sub use cases for AI/ML-based Beam Management use case.*** |
| LGE [15] | **Proposal #1: Consider DL Tx beam prediction in time domain with priority for sub-use cases of AI/ML for beam management.**  **Proposal #2: Potential specification impact for DL Tx beam prediction can be additional UE reporting which can include assistant information for gNB side AI/ML or predicted beam information from UE side AI/ML.** |
| CIACT[16] | ***Proposal 1: AI/ML based algorithm could be used to simplify the beam measurement process.***  ***Proposal 2: Both AI/ML model(s) from gNB based solution and preassemble AI/ML model(s) based solution could be considered for further study.*** |
| Apple[17] | ***Proposal 1: Study spatial domain beam prediction with measurement for limited number of beams as well as a flexible beam measurement and report framework to support dynamic activation/deactivation of beam measurement reference signal and beam report.***  ***Proposal 2: Study FR2 spatial domain beam prediction with FR1 measurements as well as CSI enhancement in FR1 to facilitate the beam prediction in FR2***  ***Proposal 3: Study time domain beam prediction based on past measurement results as well as TCI activation/indication to facilitate the beam prediction in time domain.***  ***Proposal 4: Study beam dwelling time prediction based on past measurement results as well as UE power saving schemes for beam measurement with regard to predicted beam dwelling time.***  ***Proposal 5: Since AI based beam prediction cannot provide 100% beam prediction accuracy, it is necessary to study hybrid AI based and non-AI based beam management.***  ***Proposal 6: Study how to management multiple AI processing simultaneously.*** |
| CMCC[18] | **Proposal 1: Spatial domain beam prediction can be a representative sub use case for beam management.**  **Proposal 2: The model inference procedure of spatial domain beam prediction includes P1 and P2 process.**  **Proposal 3: For spatial domain beam prediction, both model inference operated at gNB side and UE side can be studied.**  **Proposal 4: The same sort method of beam pairs is pre-defined so that gNB and UE have the same understanding of index of beam pairs.**  **Proposal 5: For model inference of spatial domain beam prediction at gNB side, CSI report framework and beam indication need further enhancement.**  **Proposal 6: For model inference of spatial domain beam prediction at UE side, CSI report framework and beam indication need further enhancement.** |
| DOCOMO[19] | **Proposal 1: Time-domain beam prediction should be studied as a sub use-case of beam management in Rel-18 AI/ML for AI.**  **Proposal 2: CSI report should be enhanced to improve the performance of time-domain beam prediction, if time-domain beam prediction is supported as sub use-case.**  **Proposal 3: Spatial-domain beam estimation should be studied as a sub use-case of beam management in Rel-18 AI/ML for AI.**  **Observation 1: Enhancements on beam selection policy in CSI reports might be potential specification impacts for spatial-domain beam estimation.** |
| Lenovo[20] | **Beam measurement and beam selection are important for initial beam assignment procedure during initial access in FR2.**  **Beam measurement and beam selection procedure is the key procedure for all beam management procedures.**  **Beam prediction at gNB/TRP side with model management-related collaboration between gNB and UE (i.e., Cat. 2) can be taken as a sub-use case for beam management in predictable trajectory scenario.**  **Beam selection from a larger number of candidate beams based on the measurement of a small number of configured beams at the UE side using AI model can be taken as another sub-use case.**  **Study UE/NW capability related signaling corresponding to AI-based beam management under correspondent network-UE collaboration levels.**  **Study how to signal AI related parameters for a beam measurement procedure.** |
| Spreadtrum[21] | ***Proposal 1: AI/ML based beam selection can be considered as one of the representative sub use cases.***  ***Proposal 2: AI/ML based beam prediction can be considered as one of the representative sub use cases.***  ***Proposal 3: For AI/ML based beam selection, training could be conducted by gNB, while inference could be conducted by UE for better performance.***  ***Proposal 4: For AL/ML based beam selection, support to configure AI model related information to UE.***  ***Proposal 5: For AL/ML based beam selection,***  ***The current CSI framework can be reused as starting point***  ***The 1/2-port CSI-RS resource and SSB can be reused as measurement resource***  ***Define new reporting quantity for beam that was not directly measured***  ***Define new UE processing capability for AI/ML based beam reporting***  ***Proposal 7: For AL/ML based beam prediction, AI model is implemented by gNB and transparent to UE.***  ***Proposal 8: For AL/ML based beam prediction, the Rel-17 TCI framework can be reused, no more enhancement is required.*** |
| TCL[22] | ***Proposal 1: The configuration of SSB beam scanning at initial access stage can be improved by ML.***  ***Proposal 2: The subsets of beams at the gNB side and UE side, can be constructed with an ML model to reduce the beam training overhead.***  ***Proposal 3: The UE position information is not necessary for predictive beam switching.***  ***Proposal 4: The predictive beam switching shall be discussed in sub use cases of inter-cell beam switching and intra-cell beam switching for latency reduction.***  ***Proposal 5: The beam failure detection performance can be enhanced by an AI/ML model based on historical beam measurements.***  ***Proposal 6: The new candidate beam qnew can be jointly determined by an ML model when beam failure occurs.*** |
| Nokia[23] | The potential use cases that can be beneficial from ML spatial domain beam prediction are:  Use wide beam measurements to predict the best refined beam(s).  Use a subset of wide/refined beam measurements to predict the best wide/refined beam(s).  Utilize the QoS metric to assist the beam prediction for improving system throughput and reducing latency.  Supervised learning ML model beam prediction can be used for finding the highest RSRP wide/refined beam(s) by reducing the beam management overhead significantly.  In order to find beam selections that increase the QoS class specific performance metrics, reinforcement learning approaches can be used for searching the beams than the ones with highest RSRP.  For RL based methods, occasional exploration may be needed to train and validate the ML model.  Support RAN1 to further consider spatial beam prediction based on both supervised learning and reinforcement learning for   * + Beam prediction for reducing beam management resource overhead and latency.   + QoS based beam prediction for improving system throughput and reducing latency.   For beam prediction based on supervised learning, the ML model consider the followings:   * + Model input: RSRP measurements of all/subset of Tx beams from gNB GoB#1, extra info can be included.   + Model output: Prediction of beam ranking or beam RSRP for all/subset of Tx beams from gNB GoB#2.   + **Offline training for the ML model generation.**   For beam prediction based on reinforcement learning, the ML model considers the following   * + Model input: Standard CSI measurements and optionally beam usage statistics and scheduling information if available   + Model output: best beam for QoS class or DRB   + Optionally explorative data collection and training   To enable the supervised learning gNB-based beam inference operation, the existing CSI measurement/report framework can be sufficient.  To enable the supervised learning UE-based beam inference operation  Additional gNB-UE collaboration and signaling may be needed for the ML model selection, model input, model validation and model fine-tuning.  Enhanced CSI measurement/report framework may be needed.  For supervised learning spatial beam prediction, support RAN1 to further study both gNB-based and UE-based beam inference, including   * + The necessary collaboration and signal exchange between gNB and UE for ML model input, model validation, model fine-tuning and model version selection.   + The necessary signaling for CSI reporting to enable ML model inference.   To enable gNB-based QoS aware beam prediction, additional signaling could be introduced so that UEs can expect transmissions from a beam that is not with the highest RSRP.  To enable flexibility of beam prediction in spatial domain, it may be required to have control over the ML model on the exact context of beam prediction in the spatial domain.  For spatial beam prediction, support RAN1 to further study UE-based beam inference, where beam predictions are performed according to the required context set by the gNB.  Spatial-temporal domain beam prediction can be used for reducing beam management resource overhead especially for mobile UE.  Both supervised learning and local online learning provide a framework to track the best beam over time with reduced measurement overhead.  Support RAN1 to further consider the spatial-temporal beam prediction with ML methods, including:   * + Further study spatial-temporal beam prediction with supervised learning method   + Further study spatial-temporal beam prediction local online learning method   + Further study the tradeoffs between the supervised learning method and the online learning method.   To enable the supervised learning or local online learning gNB-based spatial-temporal beam inference operation, the existing CSI measurement/report framework can be sufficient.  To enable the supervised learning or local online learning UE-based spatial-temporal beam inference operation  Additional gNB-UE collaboration and signaling may be needed for the ML model input, model validation, model fine-tuning, and model selection.  Enhanced CSI measurement/report framework may be needed.  Framework supporting model reliability may be needed.  For supervised learning or local online learning spatial-temporal beam prediction, support RAN1 to further study both gNB-based and UE-based beam inference, including   * + The necessary collaboration and signal exchange between gNB and UE for ML model input, model validation, model fine-tuning, model-reliability, and model version selection.   + The necessary signaling for reporting the ML model output. |
| Intel[24] | **For AI/ML assisted beam management use-case, consider both spatial and temporal domain beam prediction sub-use cases within the purview of an offline learning framework.**  **The ML model may reside either at UE or gNB**  **One possible area of specification impact for AI/ML model integration may be for triggering of beam measurement reports and reference signal transmissions, as well new L1 reporting formats.** |
| NVIDIA[25] | **Observation 1: AI/ML techniques can be used to predict beam in time and spatial domain, which can reduce the overhead and latency associated with the P1/P2/P3 beam management procedures.**  **Proposal 1: Beam prediction should be selected as one representative sub use case.**  **Proposal 2: Study the signalling support for the training and execution of AI/ML models for beam prediction.**  **Proposal 3: Study the data required by AI/ML models for beam prediction (e.g., data reported by UE to gNB, assistance data from gNB to UE).**  **Proposal 4: Study how to deliver outputs generated by AI/ML models for beam prediction from gNB to UE and from UE to gNB.** |
| AT&T[26] | **Proposal 1: Both centralized (e.g. across multiple gNBs) and decentralized (e.g. a single gNB or UE) approaches for AI/ML based beam management should be evaluated.** |
| Mavenir [27] | **Proposal 1: RAN WG1 should study the application of AI/ML-based algorithms on the following Use Cases to understand if any specification improvement is required:**  **Reducing the P-1 overhead by optimizing the intra/inter TRxP (e.g., gNB) transmit beam-sweeping directions and periodicity.**  **Reducing the P-2 overhead by predicting the best set of P-2 beams.**  **Improving the beam-based mobility by determining beam-specific cell individual offset (CFO) and Time-to-Trigger (TTT).**  **Proposal 2:** **RAN WG1 should consider the following KPIs while studying the application of AI/ML-based algorithms in beam management:**  **Throughput**  **Beam-switching success rate**  **Number of too early beam-switches**  **Number of too late beam-switches**  **Number of ping-pong cases**  **Link failure rate**  **Outage rate** |
| QC[28] | **Proposal 1: RAN1 should study temporal beam prediction and identify aspects of temporal beam prediction where AI/ML-assisted methods are beneficial.**  **Proposal 2: For UE-side training, RAN1 should focus on offline training scenario, in which the development and training of the AI model for temporal beam prediction happens offline without the need to involve 3gpp signaling.**  **Proposal 3: RAN1 should study the signalling aspects related to gNB sending assistance information to help UE with data collection for training, for the purpose of temporal beam prediction.**  **Proposal 4: RAN1 should study and evaluate the benefits of temporal beam prediction at UE and gNB and the associated signalling needed to assist or enable beam prediction at each side.**  **Proposal 5: For temporal beam prediction, RAN1 should study the signalling aspects related to exchanging information about beam prediction quality and a metric for beam prediction quality**  **Proposal 6: For temporal beam prediction, RAN1 should study the signalling aspects related to gNB sending assistance signalling to help UE in comparing predicted measurements with actual measurements.**  **Proposal 7: RAN1 should study codebook-based spatial domain beam prediction and identify aspects of codebook-based spatial domain beam prediction where AI/ML-assisted methods are beneficial.**  **Proposal 8: For UE-side training, RAN1 should focus on offline training scenario for codebook-based spatial domain beam prediction, in which the AI/ML model design and training does not involve 3gpp signalling.**  **Proposal 9: For UE-side training, RAN1 should study the signalling aspects related to gNB sending assistance information to help UE with data collection for training, for the purpose of codebook-based spatial domain beam prediction.**  **Proposal 10: RAN1 should study and evaluate the benefits of codebook-based spatial (+time) domain beam prediction at UE and gNB and the associated signalling needed to assist or enable beam prediction at each side.**  **Proposal 11: For spatial domain beam prediction, RAN1 should study the signalling aspects related to exchanging information about beam prediction quality and a metric for beam prediction quality.**  **Proposal 12: For spatial domain beam prediction, RAN1 should study the signalling aspects related to gNB sending assistance signalling to help UE in comparing predicted measurements with actual measurements.**  **Proposal 13: RAN1 should study methods for non-codebook-based spatial domain beam prediction and study signalling aspects needed to enable such a prediction.** |
| Fujitsu[29] | **Proposals 1: To limit the workload for evaluation, beam predication in spatial domain can be selected as the only sub use case for beam management.**  **Proposal 2: New signaling carrying beam information in spatial domain should be studied.** |
| Charter[30] | **Observation 1**:DFT based codebooks break down when the angular spread increases (correlation decreases); it is possible to train a ML/DL network in order to derive and update a vector-quantized codebook for beam management on the gNB side, using ML/DL.  **Proposal 1:**Consider the option to enhance beam management with a dynamic vector-quantized codebook based on SVD and ML, and have it exchanged with the UE using appropriate interaction mechanisms between gNB and UE(s). |

# Reference

1. R1-2203143 Discussion on AI/ML for beam management Huawei, HiSilicon
2. R1-2203251 Discussion on potential enhancements for AI/ML based beam management ZTE
3. R1-2203284 Discussions on AI-BM Ericsson
4. R1-2203375 Discussion for other aspects on AI/ML for beam management InterDigital, Inc.
5. R1-2203454 Discussion on other aspects on AI/ML for beam management CATT
6. R1-2203553 Other aspects on AI/ML for beam management vivo
7. R1-2203691 Discussion on other aspects on AI/ML for beam management NEC
8. R1-2203730 Consideration on AI/ML for beam management Sony
9. R1-2203811 Other aspects on AI/ML for beam management xiaomi
10. R1-2203900 Representative sub use cases for beam management Samsung
11. R1-2204018 Other aspects of AI/ML for beam management OPPO
12. R1-2204060 Beam management with AI/ML Beijing Jiaotong University
13. R1-2204078 Discussion on sub use cases of beam management Panasonic
14. R1-2204103 Discussion on sub use cases of AI/ML for beam management use case FUTUREWEI
15. R1-2204152 Other aspects on AI/ML for beam management LG Electronics
16. R1-2204183 Discussions on AI-ML for Beam management CAICT
17. R1-2204241 Enhancement on AI based Beam Management Apple
18. R1-2204298 Discussion on other aspects on AI/ML for beam management CMCC
19. R1-2204378 Discussion on other aspects on AI/ML for beam management NTT DOCOMO, INC.
20. R1-2204420 Further aspects of AI/ML for beam management Lenovo
21. R1-2204501 Discussion on other aspects on AI/ML for beam management Spreadtrum Communications
22. R1-2204569 Discussions on Sub-Use Cases in AI/ML for Beam Management TCL Communication
23. R1-2204574 Other aspects on ML for beam management Nokia, Nokia Shanghai Bell
24. R1-2204796 Use-cases and specification for beam management Intel Corporation
25. R1-2204843 On other aspects of AI and ML for beam management NVIDIA
26. R1-2204863 System performance aspects on AI/ML for beam management AT&T
27. R1-2204938 AI/ML for beam management Mavenir
28. R1-2205027 Other aspects on AIML for beam management Qualcomm Incorporated
29. R1-2205079 Sub use cases and Spec impact on AI/ML for beam management Fujitsu Limited
30. R1-2205094 Discussion on Codebook Enhancement with AI/ML Charter Communications, Inc

# Appendix: Previous Agreements

void