**3GPP TSG-RAN WG3 Meeting #112electronic R3-212689**

**Online, 17th – 27th May 2021**

**Agenda Item: 18.3**

**Source: CMCC (moderator)**

**Title: SoD Data Collection Use Cases**

**Document for: Discussion and Decision**

# 1 Introduction

**CB: # 46\_DataColl\_UseCases**

**- Chair: suggest to structure discussion around 5 areas, splitting work among companies for the resulting TPs:**

**1) Common parts / overview / general descriptions (merging any agreeable parts from e.g. 2301, 2313, 1683, 2190) (NEC?/E///?)**

**2) ES/EE (merging any agreeable parts from e.g. 2030, 2031, 2523, 2507, 1669, 2315) (ZTE,CU,Len,Moto?/CMCC?)**

**3) Load Balancing / Load Prediction (merging any agreeable parts from e.g. 2032, 2033, 2504, 2524) (CMCC?/HW?)**

**4) Mobility / Traffic Steering (merging any agreeable parts from e.g. 2028, 2029, 2191, 2271, 2269, 2316) (CATT?/ID?)**

**5) Other use cases (if any / if agreeable) (from e.g. 1969, 2179, 2389)**

(CMCC - moderator)

There are some overlaps between CB#46 and CB#47，we found it is sensible to discuss the standard impacts with the specific solutions together, so after coordination with the moderator of CB#47\_DataColl\_StdImpact, this CB#46 will only focus on use case description and potential benefits. The solutions for the identified use cases and potential standards impact will be discussed in CB#47.

Under the chair’s guidance, the offline discussion consists of 5 areas, the discussion will be structured as follows:

- In the first round of offline discussion, it is proposed to collect company views on the general question for each area;

- In the second round of offline discussion, it is proposed to focus on TPs for use cases, including ES/EE, Load balancing/ Load Prediction, Mobility / Traffic Steering, as well as common parts and other use cases if agreeable.

# 2 For the Chairman’s Notes

**To be added after email discussion.**

# 3 Discussion

## 3.1 Common

The proposals in this part (from [15] [16] [11]) are diverse and seems cannot to merge, TPs for prioritized three use cases are provided in [2] and could be discussed in following corresponding use case parts.

Companies are invited to provide views on whether following proposals are agreeable.

**Q1: Which following proposal(s) from [15] is agreeable for you?**

**Proposal 1: “Confidence Level” of a model inference output should be supported and provided to “Actor” node.**

**Proposal 2: “Validity Time” of model inference output should be supported and provided to “Actor” node.**

**Proposal 3: For ML enabled load balancing, “Model Training” can be located at CN/OAM or NG-RAN (CU). “Model Inference” can be located at NG-RAN (CU).**

**Proposal 4: For ML-enabled load balancing, input data of Model Training includes QoS requirement, traffic volume, cell capacity, cell resource status, and measurement data.**

**Proposal 5: For ML-enabled load balancing, output data of Model Inference includes prediction of traffic volume, cell capacity and cell resource status, action of UE (such as handover), etc.**

**Proposal 6: For ML-enabled load balancing, the prediction results of traffic, resource status, cell capacity, etc., should be exchanged between the source gNB CU and neighbouring gNB-CUs via Xn interface.**

**Proposal 7: For ML-enabled load balancing, handover decision at source cell can be decided by the prediction results of its own and neighbouring cells.**

**Proposal 8: NG-RAN can decide to switch off certain cells to lower energy consumption based on the prediction of cell load information.**

**Proposal 9: For ML-enabled energy saving, a gNB can send a message and avoid deactivating a needed cell of a neighbouring gNB, by utilizing the prediction result of cell load information exchanged between them.**

|  |  |  |
| --- | --- | --- |
| **Company** | **Yes/No** | **Reasons/Comments/Suggestions** |
| Futurewei | Proposal 1: No.  Proposal 2: Yes with clarification.  Proposal 3: No.  Proposal 4-5: Yes.  Proposal 6: Yes.  Proposal 7: No.  Proposal 8: No.  Proposal 9: No. | Proposal 1: Not all ML models provide confidence interval (CI). If we want CI to be provided, then first we need to define how CI should be calculated for those ML methods that don't directly generate such information. Before the definition is done, we don’t see CI is appropriate.  Proposal 2: need a clear definition of “Validity Time” if we want to use it. For example, in traffic prediction use case, the prediction result may be the traffic load for the next interval (which depends on the measurement interval in the input/output), thus, the prediction result is only good for the next interval naturally.  Proposal 3: this proposal is not necessary. Regardless of use cases, where the model training and model inference should be performed belong to vendor decision.  For proposal 4 and 5: while we agree to these two in principle, we have some comments.   1. Input data should be rich enough in order to generate output data. Proposal 4 does not cover all the needed measurements. For example, if the model suggests a UE handover, then the cell capacity of neighbouring cells need to be included as input data. It seems the relationship between Proposal 4 and 5 have not been carefully considered. 2. HO decision most likely does not based only on inference result as it will take into account other factors as well. It is cleaner to separate model inference and decision in any use case, while the 2 can reside on the same functional block.   Proposal 7: The decision of handover or not should be vendor-dependent, based on each vendor’s algorithm. We suggest rephrasing this proposal to something like : *“...the source cell can utilize the model inference results for itself and neighbouring cells together with other information to make HO decision...”*  Proposal 8: Again, the decision-making is vendor dependent. Can be rephased like in Proposal 7.  For proposal 9: two comments  1) not sure if this is necessary; if the target gNB needs to remain active, it does not need this request message; its own AI model should be able to make the decision (note the traffic information is exchanged between the source gNB and the target gNB so they know each other’s situation well).  2) while a gNB can send a message to neighbouring gNBs, whether to deactivate a particular cell on the gNB is up to the gNB's decision.  **In general, we suggest separating action/decision-making from inference results/prediction; we should not define how the decisions will be made.** |
| CATT | P1, P6, P7, P8, P9: yes  P2: no  P3–P5: cannot agree now (please see in the comment) | For P2, we don’t observe **exchanging** the “validity time” quite useful. It can be entirely up to implementation. Need further clarification on why exchanging this is necessary.  For P3–P5, in our understanding load balancing encompasses mainly three sub-functions: load prediction, UE traffic and mobility prediction, and load balancing decision.  The two “prediction” sub-functions can surely benefit from AI/ML (e.g. supervised learning), but once the “accurate” prediction result is provided, we doubt if it is really beneficial to make the load balancing decision function itself an AI-assisted one, e.g. by using reinforcement learning.  P3–P5 are mainly about taking decision, so we cannot agree with them now. |
| Samsung | P1, P2: need to study case by case  P3, P6, P7, P8: Yes  P4, P5: Need more clarification/discussion  P9: Need to discuss the exchange of predicted deactivation decision firstly | P1 and P2: Interface impact needs to be studied case by case. “Confidence level” and “validity time” can be considered for further detailed use case study.  P4 and P5: input/output highly depends on the AI functionality. Suggest to discuss the AI functionality for load balancing first, and then to discuss the related input/output for each functionality.  P9: this message to avoid deactivation is based on the exchange of predicted deactivation decision. So prefer to discuss whether to support exchange predicted deactivation decision firstly. |
| CMCC |  | Not sure P1, P2 are necessary for all AI models, and whether confidence level and validity time are useful and/or trustful.  It is better to discuss P3-P9 in the standard impact study of corresponding use case. |
| Nokia | P1: Not sure  P2, P4, P5: No  P3, P6,P7,P8, P9: Yes | P1: How can confidence level be useful to the Actor? If we consider confidence level over the predictions, this only tells us that prediction results are stable and within some range but it doesn’t tell anything about how good the prediction is (with respect to the real measurements.  P2: How can you know in advance how long an ML model will be valid?  P4, P5: Decisions on those details should be postponed to later stage.  P3: We agree though we believe it would be good not to mix CN and OAM when it comes to training. Also, those options should be understood just as possibilities and not as restrictive options. |
| NEC |  | Overall proposals 1-9 are too detailed for use case description and look like solution. While we sympathize some of the ideas in proposals 1-9, they fill too limiting for use cases.  Proposal 1 and 2: This could be useful for some outputs, but not for all. Looks like more intended for interim predictions, not for final decisions.  Proposal 3: Such conclusion should be made after analysing all provided solutions.  Proposal 4: Could be some of the inputs depending on ML model/algorithm, but not limited to. We could summarize later based on provided solutions.  Proposal 5: Could be examples of outputs depending on ML model/algorithm, but not limited to. We could summarize later based on provided solutions.  Proposal 6: Depends on solution. Not always necessary.  Proposal 7: Depends on solution. Not always necessary.  Proposal 8: It is possible. But this is not the only possible solution.  Proposal 9: It is possible. But this is not the only possible solution. |

**Q2: Which following proposal(s) from [16] is agreeable for you?**

1. It is proposed to sufficiently progress the prioritized use cases on energy saving, load balancing, traffic steering/mobility optimization, i.e. at least by identifying their impact on the specifications, before considering any new use case.
2. It is proposed to investigate standard support to enable machine learning at the network side to reduce the energy consumption at the UE, such as the introduction of new UE feedback information reflecting the UE’s energy consumption.
3. It is proposed to investigate a RAN-based Energy Efficiency solutions to achieve better energy consumption at the RAN, for example by means of exchange of this information.
   * + **Exchanging information between RAN nodes, describing if a certain RAN action is taken due to energy efficiency**
     + **Exchanging of RAN node energy-related information between RAN nodes**
     + **Exchanging performance feedback related to a certain energy efficiency action taken in another RAN node**
     + **UE traffic prediction**
     + **UE mobility prediction**
     + **UE performance feedback**
4. RAN3 to discuss the following improvements related to Traffic Steering:
   * + Improved Mobility Load Balancing decisions using UE traffic related predictions and RAN load predictions.

**Proposal 5: It is proposed to consider the Link Adaptation use case as low priority.**

|  |  |  |
| --- | --- | --- |
| **Company** | **Yes/No** | **Reasons/Comments/Suggestions** |
| Futurewei | Proposal 1: Yes  Proposals 3 and 4: Yes, with clarification  Proposal 2 and 5: No | Proposal 2: The discussions so far were for RAN power saving, not for UE. So UE power saving should be considered a new use case. As a new use case, it should have lower priority and should be postponed.  Proposal 3:  1) Item 3: To clarify: what does "performance feedback” mean?  2) Item 4: To clarify: what does UE traffic prediction mean, per UE? If yes, this would require a new use case to predict UE traffic?  3) Item 5: similar as the above, UE mobility prediction is not included in the use cases being proposed so far.  In general, we believe ML-based RAN energy saving benefit can be evaluated at OAM or other-higher level entity instead of exchanging all the information between RAN nodes.  Proposal 5: This proposal is not necessary. If other use cases score higher based on agreement, then link adaptation use case becomes lower priority naturally. |
| CATT | P1, P4, P5: yes.  P2: no.  P3: cannot agree now. | For P2, we don’t think UE power saving belongs to the agreed “prioritised use cases”.  For P3, similar with the comment in the tabular above, need to clarify the benefit of AI-assisted MLB decision, assuming that AI-assisted predictions are already supported. |
| Samsung | P1: No  P2: No  P3: Prefer to identify AI functionality first, and then to study input/output and interface impact.  P4: Yes  P5: Partly yes | P1: prefer to identify the new potential use case in parallel with studying the standard impact of agreed use cases.  P2: UE power saving seems not belong to the agreed use case. And it is more relevant to other WGs. Slightly prefer to down-prioritize it.  P3: the input/output and interface impact highly depends on ML model functionality. Suggest to identify ML model functionality firstly.  P5: it seems that link adaptation is highly relevant to other WGs. Slightly prefer to down-prioritize it. |
| CMCC | P1: Yes  P2-P4: cannot agree now | For P2, agree with CATT and Samsung that UE power saving does not belong to the agreed “prioritised use cases”.  It is better to discuss P3-P4 in the standard impact study of corresponding use case. |
| Nokia | P1:Yes  P2, P3, P4, P5: No | P2: This SI is about introducing ML in the RAN. To our understanding it is not about studying UE improvements but RAN improvements.  P3, P4: We should not focus on UE predictions and extensive new UE measurements that may have high UE impacts. This is a RAN3 led SI where minimal input can be expected from other groups, e.g., RAN2, based on LSs.  P5: We should focus on the approved use cases and reach some sufficient baseline description before studying new ones. We have already agreed on the prioritized use cases. |
| NEC |  | Proposal 1.  It is fine for us to concentrate on merging prioritized use cases:  - Network energy saving  - Network load balancing  - Mobility optimization  - Traffic steering.  For us it is also fine to define supporting use cases:  - UE position prediction  - Traffic prediction  - Load prediction.  Also, it is important to distinguish first group above and second group above. For example, Network load balancing is different from Load prediction. Mobility optimization is different from UE position prediction.  Proposal 2.  No problem for us if time allows given the priorities mentioned in our comment to Proposal 1.  Proposal 3.  This depends on solutions. Better first to see solution description, then decide which information should be exchanged.  Proposal 4.  Traffic steering is prioritized, so we support this use case. Second half of the proposal is solution. Better to see full solution description before decision.  Proposal 5.  This looks like RAN1/RAN2 topic. |
|  |  |  |
|  |  |  |

**Q3: Which following proposal(s) from [11] is agreeable for you?**

Proposal 1: RAN3 is proposed to consider splitting the use cases into two parts: “AI-assisted decision” where the output is the decision itself (i.e. to turn on/off a cell, to hand over a UE from this cell toward that cell), and the “AI-assisted prediction” where the output is merely the prediction result used as an intermediate input for decision making.

Proposal 2: If proposal 1 is agreed, RAN3 is proposed to focus firstly on “Load prediction” “UE location prediction”, “Service data prediction” and “AI-assisted handover / data offloading decision”.

Proposal 3: If proposal 1 is agreed, RAN3 is proposed to confirm that “cell/carrier/symbol/etc turning on/off decision” and “radio resource configuration / scheduling” can be affected by the aforementioned predictions.

|  |  |  |
| --- | --- | --- |
| **Company** | **Yes/No** | **Reasons/Comments/Suggestions** |
| Futurewei | Proposal 1-3: No | There is no need for these proposals. Whether to use ML output or inference to make decision directly or just as part of the decision-making is implementation dependent. |
| CATT | Slightly prefer splitting the use cases into “AI-assisted decisions” and “AI-assisted predictions” (or identically, “tool-box use cases”). | An alternative option: includes only the decisions (both AI-assisted and non-AI-assisted conventional ones) into the “use case” section, but add a new section parallel to the “use case”, maybe namely “tool box”. |
| Samsung | Slightly prefer to classify the solution as “ML-generated solution” and “ML-assisted solution” | For one specified use case, the solution can be classified based on ML model functionality, as “ML-generated solution” and “ML-assisted solution”.  For example, for use case of load balancing, one functionality of ML model is to predict load, and ML-assisted solution is load balancing strategy obtained based on the predicted load; or ML model functionality is to generate load balancing strategy, and ML-generated solution is the load balancing strategy obtained from ML model directly. |
| CMCC | Proposal 1-3: cannot agree now | Not sure P1 is necessary. |
| Nokia | P1, P2, P3: No | About P1, P2 and P3: We don’t see it necessary to split use cases in different categories. We have identified the use cases we will study. In our view “load prediction”, “UE location prediction”, “service data prediction” and “AI-assisted handover/data offloading decision” are enablers to study different use cases and can be used transparently for the identified use cases. |
| NEC |  | Proposal 1.  No problem to have two categories of use cases. However, naming could be improved to avoid confusion with prediction/classification.  First category: Output is directly used by RAN.  Second category: Output is used by other ML models as input.  Proposal 2.  Prioritized use cases are:  - Network energy saving  - Network load balancing  - Mobility optimization  - Traffic steering.  First, these use cases should be merged.  For us it is also fine to define supporting use cases:  - UE position prediction  - Traffic prediction  - Load prediction.  Proposal 3:  Such decision is better to make after analysing provided solutions to use cases. |
|  |  |  |
|  |  |  |

## 3.2 EE/ES

Several papers provide views and potential TPs on the use case of energy saving. Moderator would like to suggest companies to discuss the potential TP for EE/ES based on 2030[6], where detailed use case description, simulation results and potential benefits are provided, and merge agreeable parts in [1][2][17][21][22].

**Q4: Do you agree to work on the potential TP for use case description of EE/ES based on [6] and merge other TPs?**

|  |  |  |
| --- | --- | --- |
| **Company** | **Yes/No** | **Reasons/Comments/Suggestions** |
| Futurewei | Yes. |  |
| CATT | Yes |  |
| Samsung | Yes |  |
| CMCC | Yes |  |
| Nokia | Yes | Small clarification perhaps in the text that we don’t consider AI but rather ML (ML techniques, ML Energy Saving, ML algorithms, etc.). |
| NEC |  | Yes. We prefer “Network Energy Saving” as a name for the use case. |
|  |  |  |
|  |  |  |

## 3.3 Load Balancing/ Load Prediction

Moderator would like to suggest companies to discuss the potential TP for load balancing based on 2524[23] and merge agreeable parts in [2][4][8][20].

**Q5: Do you agree to discuss the potential TP for use case description of load balancing based on [23] and merge other TPs?**

|  |  |  |
| --- | --- | --- |
| **Company** | **Yes/No** | **Reasons/Comments/Suggestions** |
| Futurewei | Yes. |  |
| CATT | Generally yes, but | Not quite sure of the following bullet in [23]:   * The load balancing based handover highly depends on the measurement report from the UE. The UE measurement configuration and measurement reports may cause amount of signaling overhead over Uu interface, UE power consumption and data interruption of running service.   We are not sure the current MLB mechanism *does* cause “amount of signalling overhead”, so maybe we needn’t mention it here. The remaining two bullets are already sufficient enough. |
| Samsung | Partly yes | Same view with CATT. It is unclear why “measurement based HO” causes large overhead.  *-The load balancing based handover highly depends on the measurement report from the UE. The UE measurement configuration and measurement reports may cause amount of signaling overhead over Uu interface, UE power consumption and data interruption of running service.* |
| CMCC | Yes |  |
| Nokia | Yes, but | The use case description from TP in [23] is OK for us. We also have a problem with the second bullet but for an additional reason. It creates an impression that the use case will be about optimizing UE measurement configuration which should not be the case. |
| NEC |  | Yes for use case description part. Solution part should be discussed separately. |
|  |  |  |
|  |  |  |

## 3.4 Mobility/ Traffic Steering

Moderator would like to suggest companies to discuss the potential TP for mobility based on 2271[14] and merge agreeable parts in [2][12][18].

**Q6: Do you agree to discuss the potential TP for use case description of Mobility/ Traffic Steering based on [14] and merge other TPs?**

|  |  |  |
| --- | --- | --- |
| **Company** | **Yes/No** | **Reasons/Comments/Suggestions** |
| Futurewei | Yes. |  |
| CATT | Yes, but | Two comments:  1. It is OK to capture use case related to CHO/DAPS. However, maybe the CHO/DAPS part could be discussed after we have reached consensus on normal mobility.  2. UE mobility prediction and/or UE traffic prediction is not mention, so merging with other contributions seems necessary. |
| Samsung | Partly yes | Prefer to merge CHO/DAPS into mobility management, as CHO and DAPS are two mechanisms of mobility management.  Load balancing is agreed as a separate use case, so it is better to put load balancing related description in load balancing use case part.  Coverage optimization is a standalone use case. Whether to involve it to be the agreed use case can be discussed later. So prefer to not involve the coverage optimization part in mobility description. |
| CMCC | Yes |  |
| Nokia | Yes, but | In [14] coverage optimization is described. However, since coverage optimization is not in the agreed use cases, a use case description is not needed for it.  Also, DAPS could be considered at a later stage when the baseline mobility use case is sufficiently studied.  Finally, given that this is a RAN3-led SI we do not think it is feasible to introduce extensive Uu impacts. This comment is related to the following part of the TP:  “The relevant information needed from the UE may include, amongst others:   * UE radio conditions * UE location * UE trajectory * UE connectivity status * UE capabilities   Other UE contextual information” |
| NEC |  | No, we believe it is too detailed and limiting, more like a solution. Instead, we would propose to consider Mobility optimization use case in [2] R3-212316 or [18] R3-212316 as a starting point for merging. |
|  |  |  |
|  |  |  |

## 3.5 Other use cases

Some new use cases are proposed to discuss, including

1. URLLC [3]
2. PSCell change [10]
3. Network slicing, QoE optimization [15]
4. L1/L2 Beam Management Configuration and Operation Optimization，Multi-user MIMO Configuration and Optimization， Scheduler Tuning and Parameter Optimization [19]

Companies are invited to provide views on whether to study above new use cases.

**Q7: Do you agree to study above new cases?**

|  |  |  |
| --- | --- | --- |
| **Company** | **Yes/No** | **Reasons/Comments/Suggestions** |
| Futurewei | No. | 1) Suggest we postpone these topics to later stages.  2) Some use cases under 4 belong to other RAN groups; should they be handled by RAN3? |
| CATT |  | Bullet 2 seems already covered.  Neutral for the rest. |
| Samsung | Yes for URLLC | URLLC is a potential use case to maintain the stringent QoS requirements, as network variations (e.g. channel state, mobility, etc.) make it challengeable to keep performance in a stable near-perfect level. AI can help to do prediction or generate optimized policy with high resiliency to variations.  PDCP duplication optimization is one of the potential directions for AI-aided URLLC, as complexity dramatically increases due to up to 4 RLC entities now. AI can help to set accurate/adaptive configuration & selection among RLC entities to realize trade-off b/w QoS performance and resource efficiency. |
| CMCC |  | Prioritized use cases should be studied first, and then new use cases could be studied if time allowed. |
| Nokia | 1, 2: No  QoE Optimization: No  Network slicing: Yes  4: Maybe | URLLC: URLLC use case seems interesting but it is difficult to evaluate it without UE impacts (and heavy RAN2 involvement). Perhaps some more information is needed to evaluate the RAN3 aspect.  PSCell change could be considered as part of the mobility use case (not as a separate use case) but after we sufficiently study the basic mobility mechanism.  QoE optimization: RAN visible QoE is covered by Rel-17 WI objective, but remains a new topic. So, it is preferable to avoid parallel discussions in QoE WI and AIML SI.  Use cases in 4 could be interesting alternatives to the agreed L3 use cases. But it is a bit obscure to us at the moment how we can evaluate them and what will be the RAN3 impacts since they are pretty much focused in the DU. |
| NEC |  | We support the following prioritization based on the previous agreements and interest in supporting use cases from several companies.  Prioritized use cases are:  - Network energy saving  - Network load balancing  - Mobility optimization  - Traffic steering.  First, these use cases should be merged.  For us it is also fine to define supporting use cases:  - UE position prediction  - Traffic prediction  - Load prediction.  If there is time after above use cases, we are open to consider more use cases. |
|  |  |  |
|  |  |  |

# 4 Conclusion, Recommendations

To be edited, if needed**.**

# 5 Reference

1. R3-211669 Machine Learning Use Case for BS Power Saving Futurewei
2. R3-211683 TP to TR 37.817 Use case description NEC
3. R3-211969 Discussion on Use Cases for RAN Intelligence Samsung, Verizon Wireless
4. R3-212028 AI based UE Trajectory Prediction ZTE Corporation, China Unicom, Lenovo, Motorola Mobility, CMCC
5. R3-212029 Solution to AI based UE Trajectory Prediction ZTE Corporation, China Unicom, CMCC
6. R3-212030 AI based Energy Saving ZTE Corporation, China Unicom, Lenovo, Motorola Mobility
7. R3-212031 Solution to AI based Energy Saving ZTE Corporation, China Unicom
8. R3-212032 AI based Load Prediction ZTE Corporation, China Unicom, Lenovo, Motorola Mobility
9. R3-212033 Solution to AI based load prediction ZTE Corporation, China Unicom
10. R3-212179 AI based PSCell change Lenovo, Motorola Mobility
11. R3-212190 Discussion on use cases for AI in RAN CATT
12. R3-212191 Discussion on UE Location prediction CATT
13. R3-212269 Data Forwarding Optimization Use Case for AI InterDigital
14. R3-212271 Mobility Optimization Use Case for AI InterDigital
15. R3-212301 Use cases for AI/ML enabled NG-RAN Intel Corporation
16. R3-212313 Overview of AI/ML use cases Ericsson
17. R3-212315 AI/ML for energy efficiency use case discussion Ericsson
18. R3-212316 AI/ML traffic steering use case discussion Ericsson
19. R3-212389 Use Cases and Requirements for Artificial Intelligence in RAN AT&T
20. R3-212504 AI-based load balancing CMCC
21. R3-212507 AI-based energy saving CMCC
22. R3-212523 Further discussions on detailed procedure and potential spec impacts of energy saving Huawei
23. R3-212524 Further discussion on AI/ML assisted load balancing Huawei