RAN-R18-WS-crossFunc-Intel - Version 0.0.2
RAN

3GPP TSG RAN Release 18 Workshop RWS-210624

Electronic Meeting, 28th June – 2nd July 2021

Agenda Item: 4.1

Source: Email discussion moderator (Intel)

Title: Report from Email Discussion RAN-R18-WS-crossFunc-Intel

Document for: Information

1 Introduction

This document is the report from the email discussion prior to the Rel-18 Workshop of the contributions submitted by Intel to agenda item 4.3 'Cross-Functionalities for both eMBB and Non-eMBB Evolution'. As instructed by the chair [RWS-210002], the purpose of the pre workshop email discussions is to provide an opportunity for questions and answers to help better understanding of the proposals among companies.

Intel submitted the following contribution to agenda item 4.3

RWS-210368 Linear Packet Coding in L2 for RAN performance improvement, Intel

RWS-210373 AI/ML enabled RAN and NR Air Interface, Intel

RWS-210369 Further Enhancements of NR Mobility, Intel

2 General questions/comments

In addition to the 3 contributions submitted to agenda item 4.3, Intel submitted an overview of Rel-18 to agenda item 4 [RWS-210376]. General questions/comment to Intel related to cross functionalities for both eMBB and Non-eMBB evolution may be raised in this section.

2.1 Round 1 Questions

Feedback Form 1: General questions/comment related to cross functionalities for both eMBB and Non-eMBB evolution

2.2 Round 1 Answers

2.3 Round 2 Questions

2.4 Round 2 Answers

3 Linear Packet Coding in L2 for RAN performance improvement

Questions related to RWS-210368 'Linear Packet Coding in L2 for RAN performance improvement' may be raised in this section.

3.1 Round 1 Questions

Feedback Form 2: Questions related to RWS-210368 'Linear Packet Coding in L2 for RAN performance improvement'

1 – NTT DOCOMO INC.

- What is the considered range for k and n for the study? It seems that larger value of k and n would cause complexity/overhead increase but they may not provide big performance gain compared with smaller value of k and n, i.e., the performance gain of the proposed LPC is mainly based on the coding rate k/n and independent characteristics of channels/links where k coded segments are transmitted, correct? Any other drawback of larger value of k and n, e.g., latency?
- Is the proposed LPC applicable/beneficial to any scenario having multiple channels/links with low correlation?
- Does the proposed LPC first target DL, UL or both?
- Which WG(s) is responsible for this study?

2 – vivo Mobile Communication Co.

We think it is worth considering network coding. We would like to understand, if the NWC is modeled above RLC layer, what's the additional computation complexity to upper layers in comparison to existing upper layer processing?

For some cases such as XR and IIoT, the services are delay sensitive. The network coding may introduce additional processing delays. What's the additional processing delay due to the additional computation from network coding and decoding?

3 – ZTE Corporation

We also believe that NW code has benefits such as improving the reliability and throughput, reducing latency, etc.

According to our Tdoc RWS-210484, using erasure coding can also improve the performance compared with simple duplication. Do you consider other coding scheme than linear packet coding in L2? Which kind of linear packet code is applied in the simulation?

Furthermore, the packet code can be potentially used in multiple scenarios, such as PDCP duplication enhancement, IAB. Is there any difference of the placement of the packet code in these scenarios?

4 – Lenovo Mobile Com. Technology

Network coding can be used to exploit the BH link diversity for IAB. According to the Page 9 of the slides, why the intermediate IAB node1 need to have the LPC layer? Does it means the intermediate IAB node need to recode the packets?

5 – Qualcomm CDMA Technologies

Q1: what does 'applying coding per packet' mean? The figure shows L2 PDU, is it SDU or PDU?
Q2: Page 4 says 'efficient use of route diversity', how to achieve 'no dependency between coded segments transmitted over the radio'?
Q3: Page 4, how to do the 'soft combing' in the upper layer?
Q4: Page 5, is the target use case IAB network?
Q5: why only assume k =2 or 4 such small number of packet segments?
Q6: Does the 'coded segment' header include the PDCP SN?

6 – Guangdong OPPO Mobile Telecom.

Q1: if a packet is chopped into k blocks and encoded into n codewords, then what is the minimum m received blcoks so that decoding is always feasible?

3.2 Round 1 Answers

NTT Docomo

Thanks a lot for good questions! We organized our answers as follows:

(Q1) *What is the considered range for k and n for the study? It seems that larger value of k and n would cause complexity/overhead increase but they may not provide big performance gain compared with smaller value of k and n, i.e., the performance gain of the proposed LPC is mainly based on the coding rate k/n and independent characteristics of channels/links where k coded segments are transmitted, correct?*

k is number of input segments for encoding, and n is number of output coded segments after encoding. In fact, fixing k/n (code rate) and increasing k and n would improve the reliability performance as long as the generated coded segments go through independent erasures. However, the simulation results shown in Appendix have dependency constraints (i.e. all coded segments in the same TB are erased together), in which case increasing k beyond a certain value (determined by the constraints) would not bring more benefits.

But even very small values of k (with less complexity and overhead) have potentials to bring huge performance gain as long as coded segments go through independent erasures, which we are targeting.

(Q2) *Any other drawback of larger value of k and n, e.g., latency?*

Yes, latency can be one drawback. And since a header should be prepended for each coded segment, if we assume the same packet size, an increased overhead ratio can be another drawback. Encoding/decoding complexity could also increase with large value of k or n.

(Q3) *Is the proposed LPC applicable/beneficial to any scenario having multiple channels/links with low correlation?*

Yes, that is the correct understanding.

(Q4) *Does the proposed LPC first target DL, UL or both?*

We are targeting both UL and DL.

(Q5) *Which WG(s) is responsible for this study?*

We think RAN2 should be the main responsible WG, given that the main impacts are on Layer-2.

Vivo

Thanks a lot for your considerations and good questions! We organized our answers as follows:

(Q1) *We think it is worth considering network coding. We would like to understand, if the NWC is modeled above RLC layer, what's the additional computation complexity to upper layers in comparison to existing upper layer processing?*

Yes, Linear Network Coding is one of candidate coding types for our LPC (i.e. if applied per packet basis).

Additional computation complexity comes from encoding and decoding operations. Encoding is by linear combination of k input segments, and decoding is by gaussian elimination (see slide 6 for our complexity analysis for more details). Those computations are all by series of addition or multiplication over finite field. Addition is at low cost as it can be done by bit-wise addition modulo 2. Direct multiplication could cost high, but we think we can circumvent such high cost direct computation by a table look-up storing all the multiplication results of finite field symbols. Such table look-up can be used for computing addition as well. And with a 1-byte finite field size, the size of such table takes only 64 KB (or even less with some compressions).

(Q2) *For some cases such as XR and IIoT, the services are delay sensitive. The network coding may introduce additional processing delays. What's the additional processing delay due to the additional computation from network coding and decoding?*

Additional processing delays purely comes from additional encoding and decoding operations that requires to perform a series of addition or multiplication over finite field. But as we mentioned, we think that, if the table look-up method can be used, additional delay would be low.

Since our LPC is aiming to apply coding "per packet basis", coding decision can be made per packet and can be chosen not to apply coding for some packets, if processing delay or computation becomes a concern.

But one point we would like to emphasize is that LPC incurs lower delay than ARQ (i.e. retransmission). If we want to enhance reliability anyway, we think LPC is a better choice if low delay is desired.

ZTE

Thanks a lot for your support and good questions! We organized our answers as follows:

(Q1) *According to our Tdoc RWS-210484, using erasure coding can also improve the performance compared with simple duplication. Do you consider other coding scheme than linear packet coding in L2? Which kind of linear packet code is applied in the simulation?*

We think it is better to be the codes satisfying MDS (maximum distance separable) property, but all linear codes are OK and up for study.

For simulations, we used the linear network coding way of encoding and decoding (see slide 10 in Appendix for more details). But the codes themselves are not specific – they can be any linear codes that satisfy MDS property, which we are targeting.

(Q2) *The packet code can be potentially used in multiple scenarios, such as PDCP duplication enhancement, IAB. Is there any difference of the placement of the packet code in these scenarios?*

If applied to URLLC scenario where PDCP duplication is applicable, we think the placement is better to be below PDCP (but also possible to be merged into PDCP layer, which is up for study). If applied to IAB, the placement can be above BAP (for LPC within IAB network, transparent to UEs) or can be below PDCP as the same for URLLC scenarios (so transparent to IAB nodes). Details can be found in slide 9 in Appendix.

Lenovo

Thanks a lot for the good question! We organized our answer as follows:

(Q) *Network coding can be used to exploit the BH link diversity for IAB. According to the Page 9 of the slides, why the intermediate IAB node1 need to have the LPC layer? Does it means the intermediate IAB node need to recode the packets?*

The intermediate IAB node does not always need to re-encode the packets to achieve gains end to end. Though we are not currently considering this re-encoding within the intermediate NW, but when needed, e.g., one hop of the path becomes less reliable, the intermediate node can re-encode to enhance the reliability with more redundancies only at the local level, without unnecessarily overload the whole path, which is up for study.

Qualcomm

Thanks a lot for good questions! We organized our answers as follows:

(Q1) *What does 'applying coding per packet mean? The figure shows L2 PDU, is it SDU or PDU?*

While in slide 3 we showed L2 PDU as the original packet for comparison with Rel-16 packet duplication (which duplicates PDCP PDU as the simplest "repetition" fashion), any SDU or PDU in L2 is up for study.

The meaning of "coding per packet" is that we are not aiming to apply coding across multiple packets as to what the traditional Network Coding scheme aimed for. In this case, we think latency can be an issue, which is not suitable to 3GPP (most real-life packets are delay-constrained), and reliability gain may not always lead to

help the overall system. But coding "per packet basis" can always guarantee per packet reliability gain to be converted into overall system throughput increase.

Moreover, "coding per packet" allows flexibility to apply coding or not per packet basis based on various considerations, e.g. resource usage, latency, or delivery status, etc.

(Q2) *Page 4 says 'efficient use of route diversity', how to achieve 'no dependency between coded segments transmitted over the radio'?*

Encoding k input segments into n output coded segments, we can decode the original packet as long as we receive "any" m segments ($m \geq k$) out of n (see slide 10 in Appendix for more details). If we use the linear block codes satisfying MDS (maximum distance separable) property then, $m = k$ is sufficient.

This makes sure no dependency when output coded segments are distributed and transmitted over multiple routes. This also allows us to efficiently use route diversity by adaptively choosing how many coded segments to transmit over each route based on its available capacity changing over time.

On the other hand, if we consider Rel-16 packet duplication, a PDCP PDU duplicated and transmitted over multiple routes are treated independently to restore the PDCP PDU. Each PDCP PDU may be split and transmitted due to RLC segmentation over each route, but those segments have dependency – i.e. different segments transmitted over different routes cannot be combined to restore the PDCP PDU.

(Q3) *Page 4, how to do the 'soft combing in the upper layer?*

What we meant by soft combining in L2 is not the bit-level soft combining done in PHY FEC. As there is no dependency between coded segments transmitted over the radio, even if in the initial transmission enough coded segments (m) were not received to restore the original packet, we can send a few more coded segments without retransmitting the original packet as a whole. This is what we meant by soft combining in L2, as those later transmitted can be used together with what has been already received to restore the original packet.

(Q4) *Page 5, is the target use case IAB network?*

IAB is one of multi-route multi-connectivity scenarios that we think LPC can be applied to bring benefits (see slide 9 in Appendix for more details). But IAB is not the only applicable scenario. Any URLLC scenarios where Rel-16 packet duplication is applicable can also be applicable.

(Q5) *Why only assume k =2 or 4 such small number of packet segments?*

There are a couple of reasons. One is that we don't want to divide the packet into segments that are smaller than a TB, since the coded segments within the same TB have dependent erasures during transmission. In this case we have no extra performance gain from increasing k, only with increased overhead. Another reason is that even very small values of k (with less complexity and overhead) have potentials to bring huge performance gain as long as coded segments go through independent erasures.

(Q6) *Does the 'coded segment header include the PDCP SN?*

The 'coded segment' header needs to include a field identifying an original packet. If a PDCP PDU is used as an original packet for LPC, then it would be natural that PDCP SN is used for the field.

Guangdong OPPO

Thanks a lot for the good question! We organized our answer as follows:

(Q1) *If a packet is chopped into k blocks and encoded into n codewords, then what is the minimum m received blocks so that decoding is always feasible?*

The minimum number of coded blocks required for decoding (m) depends on the type of code. We are considering to use any linear block code satisfying MDS (maximum distance separable) property and if MDS codes are used, m = k.

3.3 Round 2 Questions

Feedback Form 3: Round 2 questions related to RWS-210368 'Linear Packet Coding in L2 for RAN performance improvement'

3.4 Round 2 Answers

4 AI/ML enabled RAN and NR Air Interface

Questions related to RWS-210373 'AI/ML enabled RAN and NR Air Interface' may be raised in this section.

4.1 Round 1 Questions

The document is structured in two sections, one on 'AI/ML for RAN' and one on "AI/ML for Air Interface". A separate feedback form is provided for each section.

Feedback Form 4: Questions related to the section of the document on AI/ML for RAN

1 – CAICT

In page 5, distributed/federated learning is proposed in NG-RAN. Are there any typical use cases for distributed/federated learning?

2 – CATT

Thank you for the contribution. Two general questions from our side:

Q1:What's Intel's views on initial enhancement and advanced enhancement? Should we study both in Rel-18?

Q2: Model transfer between RAN and UE: Since the model is transferred between NG-RAN node and UE and it is difficult to transfer the model via RRC, the way to transfer the AI model in Uu interface needs further study.

3 – LG Electronics France

Q) Do you think it is beneficial to study both network-centric ML approach and UE-centric ML approach or it is better to only consider network-centric ML approach in Rel-18?

4 – Motorola Mobility UK Ltd.

Thank you for the interesting contribution. My question is related to your proposal in the ***"Example use-cases - MIMO"*** slide, which reads *"event-based CSI reporting"*. Can you please clarify how UL resources would be allocated for event-based CSI reporting? Also, is the event-based CSI reporting applicable to non-AI/ML approaches as well?

5 – ROBERT BOSCH GmbH

[Bosch] We feel the slides discusses more NW-centric AI/ML approaches; however, for UE-centric AI/ML approaches:

- What enhancements are needed for AI-based mobility?
- What enhancements are needed to utilize pQoS (e.g., communication service availability or reliability)?
- What enhancements are needed for AI-based positioning (assuming an vehicle-based UE with an AI/ML sensor fusion) ?

6 – Sony Corporation

Thanks for the contribution. We have a question.

- In step 3&4, ML inference model is deployed to the UE over the air interface. Does this mean that network configure the ML inference model to UE?

7 – MediaTek Inc.

Thanks for the contribution. Some questions for clarification:

1) Does Intel have any priority among the use cases mentioned?

2) Among the 4 levels of AI/ML mentioned, which level does Intel want to focus?

3) Does the 4th level (both network and the UE have the capability of training and inference) similar to Federated learning?

4) For Federated learning, does Intel have any details specific to model splitting, traffic overhead and federated averaging and convergence?

8 – NEC Corporation

NEC supports having WI in RAN3 as continuation of the current RAN3 SI and having a new SI as continuation of the current RAN3 SI covering wider scope and use cases.

NEC supports having SI on AI/ML for physical layer.

Feedback Form 5: Questions related to the section of the document on AI/ML for the air interface

1 – Qualcomm Incorporated

Are Steps 1-4 just categorization of different AI/ML types, or is the proposal to study them sequentially?

4.2 Round 1 Answers

1 – CAICT

In page 5, distributed/federated learning is proposed in NG-RAN. Are there any typical use cases for distributed/federated learning?

Answer: Thanks for the question. Currently RAN3 discussed AI/ML use cases related to load balancing, energy saving and mobility enhancement. It is noticed that, similar use cases are discussed in SA5 as well, above use cases can be first studied in Rel-18 for federated/distributed learning between NG-RAN nodes (e.g. between CN/OAM and RAN).

2 – CATT

Thank you for the contribution. Two general questions from our side:

Q1:What's Intel's views on initial enhancement and advanced enhancement? Should we study both in Rel-18?

Answer: Thanks for the questions. We think the "performance" bar for specifying advanced enhancements should be higher than that for enabling initial enhancements. Our view is to focus on initial enhancements in Rel-18.

Q2: Model transfer between RAN and UE: Since the model is transferred between NG-RAN node and UE and it is difficult to transfer the model via RRC, the way to transfer the AI model in Uu interface needs further study.

Answer: Thanks for the comment. yes, this has to be studied and possibly leveraging outcome of SA1 SID "AI/ML model transfer in 5GS"

3 – LG Electronics France

Q) Do you think it is beneficial to study both network-centric ML approach and UE-centric ML approach or it is better to only consider network-centric ML approach in Rel-18?

Answer: We are not clear about the definition of UE centric approach. Enhanced UE reporting based on AI/ML based prediction at the UE can definitely be studied. However, currently most of the PHY decisions are finally taken by the NW based on UE reporting – is this aligned with the principle of UE centric ML approach ?

4 – Motorola Mobility UK Ltd.

Thank you for the interesting contribution. My question is related to your proposal in the ***"Example usecases - MIMO"*** *slide, which reads "event-based CSI reporting". Can you please clarify how UL resources would be*

allocated for event-based CSI reporting? Also, is the event-based CSI reporting applicable to non-AI/ML approaches as well?

Answer: Thanks for the question. The exact method of how the report would be carried will have to be studied based on the criticality of the report (e.g. L3 measurements or beam failure reports etc.). Yes, AI/ML could be the motivation but not a requirement for event-based reporting – the method is ultimately up to implementation.

5 – ROBERT BOSCH GmbH

We feel the slides discusses more NW-centric AI/ML approaches; however, for UE-centric AI/ML approaches:

- What enhancements are needed for AI-based mobility?

Answer: Thanks for the question. It is not clear to us how to understand UE-centric AI/ML approach for mobility and what is the role of the network. So far, mobility handover is decided or being assisted at the network side (network also configure CHO candidates to UE). We are open to study what is the UE-centric ML approach and the level of UE flexibility, etc. Moreover, mobility enhancement is now studied as a use case in RAN3 for Rel-17. Comparison between network-based (RAN3 solution) and UE-centric approach would be helpful.

- What enhancements are needed to utilize pQoS (e.g., communication service availability or reliability)?

Answer: Thanks for the question. We are not clear about how to understand "utilize pQoS". Does it refer to traffic/service awareness at RAN side?

- What enhancements are needed for AI-based positioning (assuming an vehicle-based UE with an AI/ML sensor fusion) ?

Answer: Thanks for the question. We are open to discuss this, but we are not so sure if it really requires additional specification work. Could you clarify what do you mean by UE-centric approaches and what specification efforts you foresee for sensor fusion?

6 – Sony Corporation

In step 3&4, ML inference model is deployed to the UE over the air interface. Does this mean that network configure the ML inference model to UE?

Answer: Thanks for the question. yes, for steps 3&4, the ML model structure, weights, biases etc. should be understood by both the gNB and the UE.

7 – MediaTek Inc.

Thanks for the contribution. Some questions for clarification:

Does Intel have any priority among the use cases mentioned?

Among the 4 levels of AI/ML mentioned, which level does Intel want to focus?

Does the 4th level (both network and the UE have the capability of training and inference) similar to Federated learning?

For Federated learning, does Intel have any details specific to model splitting, traffic overhead and federated averaging and convergence?

Answer: Thanks for the questions.

1) Generally we would like to start with a small set of use-cases – may be 2-3 depending on the use-cases and TU. but we have not decided on priority yet.

2) We think the "performance" bar for specifying advanced enhancements (steps 3&4) should be higher than that for enabling initial enhancements (steps 2&3). Our view is to focus on initial enhancements in Rel-18.

3) yes

4) our view is motivated by such issues that we should develop evaluation methodology first and gradually progress towards specifying more sophisticated AI/ML implementation

8 – NEC Corporation

NEC supports having WI in RAN3 as continuation of the current RAN3 SI and having a new SI as continuation of the current RAN3 SI covering wider scope and use cases.

NEC supports having SI on AI/ML for physical layer.

Answer: Thank you for sharing your views

9 – Qualcomm Incorporated

Are Steps 1-4 just categorization of different AI/ML types, or is the proposal to study them sequentially?

Answer: Thanks for the question. We think the "performance" bar for specifying advanced enhancements (steps 3&4) should be higher than that for enabling initial enhancements (steps 2&3). Our view is to focus on initial enhancements in Rel-18.

4.3 Round 2 Questions

Feedback Form 6: Round 2 questions related to RWS-210373 'AI/ML enabled RAN and NR Air Interface

1 – Samsung Electronics Polska

Q1. In slide 7, regarding AI/ML for Air Interface, in step 2, can you clarify more about what kind of information is included in training/inference data?

Q2. In slide 8, regarding initial enhancements, for example of dynamic adapted RS training, Intel's intension is to rely on implementation based at least for AI part but the figure looks some RAN1,2 specfication impact in for optimization. Can you explain more about what is required in addition to current specification?

Q3. In slide 10, regarding CSI-compression, what kind of training mechanism do Intel assume? e.g., which data set will be used? offline or online training?

Q4: which use case require FL or which use case can benefit from FL?

Q5: In slide 5, is it feasible to share model and data between gNB even if there is cell-specific information? (e.g., building blockage)

Q6: For NN-based positioning, if NN is applied at UE, what will be feed back to gNB for positioning? Whether the report prodecure have impact on performance of positioning?

2 – MediaTek Inc.

Thanks for the answers. Just following the answers,

1) For Federated learning do you consider all UEs or a fraction of all UEs (e.g. UE's current Active UEs)?

2) Does Intel think the AI-ML model training/inference is also distributed (split) between UE and network and updated at times, or the training is centralized and only the inference is distributed?

3 – ROBERT BOSCH GmbH

Thank a lot for your answer. Our question about pQoS is a general question, which tries to understand if pQoS (as such) can to be studied in an AI/ML 3GPP framework. Therefore, our discussions assumes that:

- pQoS is put in an AI/ML framework (e.g., intel AI framework)
- in this case, the QoS needs to be predicted for, e.g., short or long term.
- Let us assume that this functionality is used to proactively enhances mobility, reliability, etc.

Hence, based on these assumptions:

- does predicting QoS require more enhancements to AI/ML frame work (NW-centric)? if yes, what are these enhancements ?
- does predicting QoS requires additionally enhancements on the UE side (assuming NW-centric as well)?
- Which groups are then impacted ?

4 – HUAWEI TECHNOLOGIES Co. Ltd.

Thank you very much for your contribution. We share similar view that we may start with initial enhancement in Rel-18, thus the study of inputs/outputs of AI model is very important.

4.4 Round 2 Answers

Response to Samsung

Q1. In slide 7, regarding AI/ML for Air Interface, in step 2, can you clarify more about what kind of information is included in training/inference data?

Answer: Thanks – yes this is a little unclear in the slides. Step-2 in slide-7 is an example where model training and inference entirely occurs at the UE side and this operation mostly can be left to UE implementation. We may expect small changes (if any) in the air-interface as a result of such operation. The training and inference data could simply be channel estimation obtained from existing RSs.

Q2. In slide 8, regarding initial enhancements, for example of dynamic adapted RS training, Intel's intension is to rely on implementation based at least for AI part but the figure looks some RAN1,2 specfication impact in for optimization. Can you explain more about what is required in addition to current specification?

Answer: Thank you. Yes, our thinking is that AI or non-AI based implementation can be left to UE choice. The additional specification impact could be dynamic indication of RS patterns or some additional reporting from UE side to enable such adaptation that can be further studied

Q3. In slide 10, regarding CSI-compression, what kind of training mechanism do Intel assume? e.g., which data set will be used? offline or online training?

Answer: Agree, data-set is a challenge. We don't have a strong view – we are open to a commonly agreed data-set (simulated channel model and/or real channel data). If it is not affecting specifications, then either online or offline training is fine – for example, the initial enhancements can use either online or offline training. For auto-encoder example in slide 10 we are assuming offline training.

Q4: which use case require FL or which use case can benefit from FL?

Answer: Thanks for the question. For RAN3 part, we are open to study and find out which use case would benef from FL. For air interface, our view is that we should develop evaluation methodology first and gradually progress towards specifying more sophisticated AI/ML implementation.

Q5: In slide 5, is it feasible to share model and data between gNB even if there is cell-specific information? (e.g., building blockage)

Answer: Thanks for the question. Sharing models and data between gNBs would be helpful to train better AI/ML model with extensive information from itself, as well as other gNBs. Even for some cell-specific information, by knowing other gNB's cell-specific information, the AI/ML model can be used to optimize inter-cell, inter-gNB performance.

Q6: For NN-based positioning, if NN is applied at UE, what will be feed back to gNB for positioning? Whether the report prodecure have impact on performance of positioning?

Ans: If NN is applied at UE for UE-based positioning then feedback should be an estimate of UE coordinate / timestamp and optionally a metric characterizing quality of coordinate estimate. If NN is applied at UE for UE-assisted positioning then feedback type may depends on how NN is utilized. There are quite many options that require further discussion and analysis for positioning. One basic example discussed in R17 is a link type (LOS/NLOS) and associated probability. Certainly it may go beyond that and relate to UE measurements as well and their quality/reliability. These aspects will require more deep analysis in 3GPP.

Response to MediaTek

1) For Federated learning do you consider all UEs or a fraction of all UEs (e.g. UE's current Active UEs)?

Answer: this is a good question that needs to be considered, we are not proposing federated learning for air-interface use cases – this is shown in slide 7 (step 4) as a possibility for future.

2) Does Intel think the AI-ML model training/inference is also distributed (split) between UE and network and updated at times, or the training is centralized and only the inference is distributed?

Answer: Both of these are possible (as shown in steps 3&4 in slide 7) but we need more study to know which use-cases can benefit and how much it can benefit from step 4 over step 3. We are not proposing steps 3, 4 for air-interface use-cases at this stage.

Response to Bosch

Thanks for the explanation and questions. In general, we think predicting QoS is an important use case for AI/ML. RAN-visible QoE is now under study in Rel-17, where RAN may also leverage RAN-visible QoE information for resource allocation and performance optimization. As we proposed on page 4, leveraging AI/ML prediction to QoS/QoE would be helpful to reduce service interruption.

For 1), In general, QoS prediction result may be used as input for other AI/ML use cases, such as load balancing, scheduling, etc. However, we don't see there's a need to enhance current AI/ML functional framework defined in RAN3 Rel-17 SI.

For 2), Additional enhancement for data collection maybe needed (from network or from UE), while we think SON/MDT framework can be used as baseline.

For 3), from our understanding, RAN2 and RAN3 would be impacted, coordination with other groups may be needed.

Response to Huawei

Answer: Thanks for sharing your views.

5 Further Enhancements of NR Mobility

Questions related to RWS-210369 'Further Enhancements of NR Mobility' may be raised in this section.

5.1 Round 1 Questions

Feedback Form 7: Questions related to RWS-210369 'Further Enhancements of NR Mobility'

1 – MediaTek Inc. **Q1.** Interruption reduction for cell group change: Do you also consider fast switching for PCells?
2 – LG Electronics France Q1 Regarding support of DAPS + (EN-DC, DC, CA, CHO and possibly FR2), could you explain more what kind of major enhancements are necessary compared to the Rel-17 DAPS and your estimation of the workload on that?

3 – vivo Mobile Communication Co.

Thanks for the good contribution. We are supportive on this working area in Rel-18, with several questions:

1. for the 1st subitem in 1st objective, I think the intention would restrict only one cell group activated at a time. In this way, the corresponding description should be added in the objective.

2. for the 2nd subitem in 1st objective, we are not sure is there use case for DAPS+MR-DC, which will be a challenge for UE.

3. Regarding the UAV part, we wonder whether the intention is to reuse the solution from LTE, or any enhancement could be considered from Intel point of view.

4. Regarding the last bullet, I am not sure whether need to restrict it to NTN use case.

4 – ZTE Corporation

On P3, it is proposed to specify the solutions to support simultaneous UL PUSCH transmission for user data. We share the same view that simultaneous PUSCH transmission during DAPS can improve UL data rate, which is important for some services. In Rel-16 DAPS, simultaneous transmission cannot be supported in some scenarios mainly due to UE capability. For example, the UE does not support power sharing in inter-frequency DAPS handover. In intra-frequency handover, RAN4 concluded that it is difficult to perform simultaneous UL transmission. So could you please clarify a bit more on this, for example, the scenarios, the solutions, etc. Thanks.

5 – HuaWei Technologies Co.

We share similar views that dynamic switch via L1/L2 signaling to support inter-cell mobility can improve the performance as in RWS-210449. In your proposal, do you think the principle of fast cell group switching is not only for MR-DC, but also for multiple carriers cases like CA, or standalone case with deployment of multiple carriers?
Do you think DAPS with CA/DC would require even stronger UE capability, i.e. supporting 4 legs at the same time?

6 – ZTE Corporation

For the fast cell group switching, will DRB level reconfiguration be allowed during the fast cell group switching? Or, only the change of cell group level configuration is allowed?

7 – ZTE Corporation

For the DAPS on DC/EN-DC, whether the UE should maintain the connection with source MN+SN and target MN+SN simultaneously?

For the mobility enhancements for NPN, will there be any impacts to CT1 and SA2?

8 – Nokia Corporation

Q1: For fast cell switching, is the assumption UE stores RRC configuration for multiple cells and RRC or L1/L2 is used to command the UE to switch from one cell to the other? I.e. is this the same as the Rel-17 objective for L1/L2 mobility?

Q2: Is the fast cell switching supposed to (only) work in FR2-FR2 case, while for other scenarios DAPS is preferred?

Q3: Which combination for DAPS is most needed in your opinion, if down-selection is to be made? DAPS+EN-DC, DAPS+DC, DAPS+CA or DAPS+CHO?

Q4: Shouldn't the NPN-specific and UAV-specific mobility objectives become a part of NPN and UAV Work Items in Rel-18 and not be handled in a generic mobility enhancements WI?

Q5: Do you have in mind solutions like mabe-before-break or rach-less for fast cell switching?

Q6: How many cell groups can be configured for UE, and how many can be active at the same time?

Q7: What would be the content of delta signalling between different releases? Does this envision "UMTS-like" critical extensions to RRCReconfiguration, or something else?

5.2 Round 1 Answers

1 – MediaTek Inc.

Q1. Interruption reduction for cell group change: Do you also consider fast switching for PCells?

Answer: Yes, we also consider fast switching for MCG<-> MCG, i.e. PCells. For some reason, it was missing in our slides.

2 – LG Electronics France

Q1 Regarding support of DAPS + (EN-DC, DC, CA, CHO and possibly FR2), could you explain more what kind of major enhancements are necessary compared to the Rel-17 DAPS and your estimation of the workload on that?

Answer: For Rel-16 DAPS, only source PCell+ target PCell is supported. To support more cells (e.g. EN-DC, DC, CA), the main work is how to handle the SCell/SCG configuration in source and target node, and corresponding capability handling. For DAPS+CHO, the main work should be how to handle failure case.

3 – vivo Mobile Communication Co.

Thanks for the good contribution. We are supportive on this working area in Rel-18, with several questions:

for the 1st subitem in 1st objective, I think the intention would restrict only one cell group activated ata time. In this way, the corresponding description should be added in the objective.

Answer: Yes, that is our intention. The final objective should be polished.

for the 2nd subitem in 1st objective, we are not sure is there use case for DAPS+MR-DC, which will be a challenge for UE.

Answer: We heard the use case from operator considering so far EN-DC has been deployed widely.

Regarding the UAV part, we wonder whether the intention is to reuse the solution from LTE, or any enhancement could be considered from Intel point of view.

Answer: Yes, our intention is to reuse the solution from LTE as much as possible.

Regarding the last bullet, I am not sure whether need to restrict it to NTN use case.

Answer: The motivation is from NTN use case. But the change should be applicable for all scenarios, i.e. up to network configuration.

4 – HuaWei Technologies Co.

We share similar views that dynamic switch via L1/L2 signaling to support inter-cell mobility can improve theperformanceasinRWS-210449. In your proposal, do you think the principle of fast cell group switching is not only for MR-DC, but also for multiple carriers cases like CA, or standalone case with deployment of multiple carriers?

Answer: Yes, we believe fast cell/cell group switching can be applied for multiple carrier scenarios, like CA, MR-DC or standalone case.

Do you think DAPS with CA/DC would require even stronger UE capability, i.e. supporting 4 legs at the same time?

Answer: Depends on the scenario. For CA case, we assume 2 legs should be sufficient, just like DC. For DC case, it will require stronger UE capability, e.g. 3 legs.

5 – ZTE Corporation

On P3, it is proposed to specify the solutions to support simultaneous UL PUSCH transmission for user data. We share the same view that simultaneous PUSCH transmission during DAPS can improve UL data rate, which is important for some services. In Rel-16 DAPS, simultaneous transmission cannot be supported in some scenarios mainly due to UE capability. For example, the UE does not support power sharing in inter-frequency DAPS handover. In intra-frequency handover, RAN4concludedthatitisdifficult to perform simultaneous UL transmission. So could you please clarify a bit more on this, for example, the scenarios, the solutions, etc. Thanks.

Answer: To our understanding, we already can support simultaneously UL transmission for inter-frequency, for ,e.g. UL feedback to source and new user data to target. Our intention is to remove the restriction. Regarding the possible limitation on intra-Freq, inter-Freq, sync, async, further study is needed in RAN1, 4.

For the fast cell group switching, will DRB level reconfiguration be allowed during the fast cell group switching? Or, only the change of cell group level configuration is allowed?

Answer: Our current thinking is, the first step should be to enable the change of cell group level configurations. It is difficult to support fast bi-direction switching if DRB level reconfiguration is allowed, especially if NAS is involved. But we are open on this.

For the DAPS on DC/EN-DC, whether the UE should maintain the connection with source MN+SN and target MN+SN simultaneously?

Answer: It depends on what capability the UE can support. If we only consider Source MN+ target MN/SN, or source MN/SN, target MN, then only 3 legs are needed. Otherwise 4 legs are needed.

For the mobility enhancements for NPN, will there be any impacts to CT1 and SA2?

Answer: We would like to limit the impact to RAN, i.e. without impact CT1/SA2 work.

8 – Nokia Corporation

Q1: For fast cell switching, is the assumption UE stores RRC configuration for multiple cells and RRC or L1/L2 is used to command the UE to switch from one cell to the other? I.e. is this the same as the Rel-17 objective for L1/L2 mobility?

Answer: There are similarities. Our intention is to have a solution for all mobility scenarios including inter-DU case .

Q2: Is the fast cell switching supposed to (only) work in FR2-FR2 case, while for other scenarios DAPS is preferred?

Answer: We would prefer to use fast cell switching not only for TRP, but also for cell and not limit to FR2-FR2 case.

Q3: Which combination for DAPS is most needed in your opinion, if down-selection is to be made? DAPS+EN-DC, DAPS+DC, DAPS+CA or DAPS+CHO?

Answer: our priority (from top to down) is DAPS+CA, DAPS+CHO, DAPS+MR-DC; FFS on DAPS+EN-DC. We are open to discuss the priority.

Q4: Shouldn't the NPN-specific and UAV-specific mobility objectives become a part of NPN and UAV Work Items in Rel-18 and not be handled in a generic mobility enhancements WI?

Answer: We think the solutions may not be limited to these specific use cases. So we think a generic WI is better.

Q5: Do you have in mind solutions like make-before-break or rach-less for fast cell switching?

Answer: Yes, we think fast cell group switching is similar to MBB, i.e. one cell group activated at a time.

Q6: How many cell groups can be configured for UE, and how many can be active at the same time?

Answer: It is related to UE capability. We would like to at least limit the maximum number of activated cell groups simultaneously, e.g. 1 or 2, in order to reduce the complexity.

Q7: What would be the content of delta signalling between different releases? Does this envision "UMTS like" critical extensions to RRCReconfiguration, or something else?

Answer: Our current thinking is to allow source to release the additional configuration as part of the HO command. But we are open to other solutions.

5.3 Round 2 Questions

Feedback Form 8: Round 2 questions related to RWS-210369 'Further Enhancements of NR Mobility'

1 – ZTE Corporation

Thanks for the clarifications provided during round 1.

For the "fast cell group switching", we also share the view that the fast change of cell group level configurations should be considered as high priority scenarios.

For the DAPS on DC/EN-DC, it seems the intention is to support three or four active legs in the DAPS HO. Since only two active connections can be supported in current specs, we want to clarify whether the multiple connectivity with more than two active connections is some kind of prerequisite of this feature?

We are open to discuss this but we think the complexity should be taken into account to ensure the feature can be implemented and deployed in the market.

5.4 Round 2 Answers

Response to ZTE

For DAPS+DC/EN-DC, simultaneous transmission/reception on three or four legs are needed unless TDM based approach or fast cell group switching is used. We do agree that the complexity should be taken into account, and therefore we are open on the candidate solutions.