**3GPP TSG-RAN WG2 Meeting #130** **R2-2503778**

**St. Julians, Malta, May 19th – 23rd, 2025**

Agenda Item: 8.3.1

Source: Mediatek Inc.

Title: Report of [POST129bis][020][AI Mob] Sim. Results Figures (Mediatek)

Document for: Discussion, Decision

# Introduction

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

* [POST129bis][020][AI Mob] Sim. Results figures (Mediatek)

Intended outcome: email discussion to gather specific comments related to result illustration

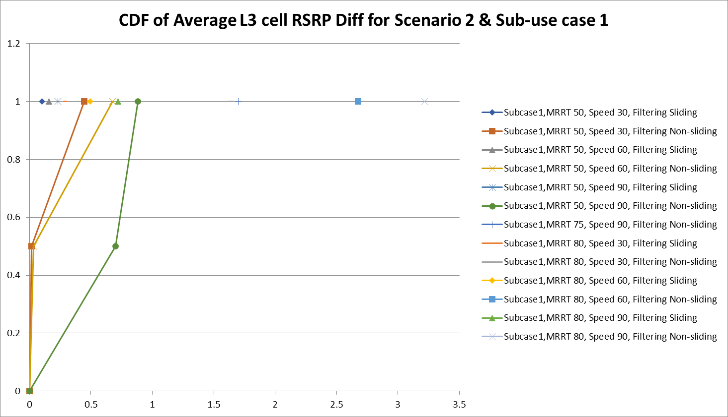
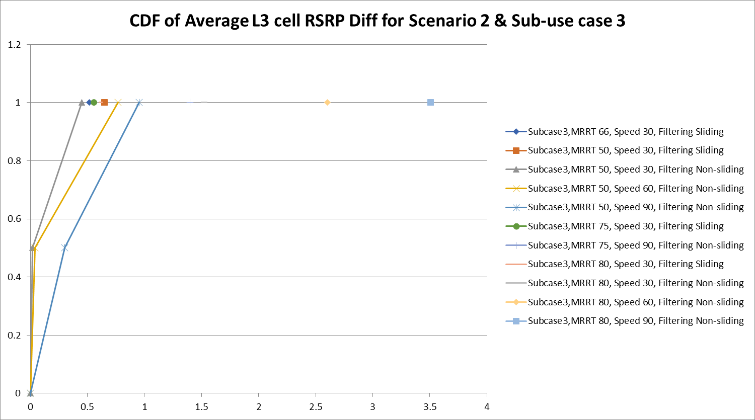
Deadline: long

The deadline for providing comments is May 2nd, 2025, at 10:00 UTC. This will allow the rapporteur sufficient time to modify the figures based on the companies' comments and suggestions.

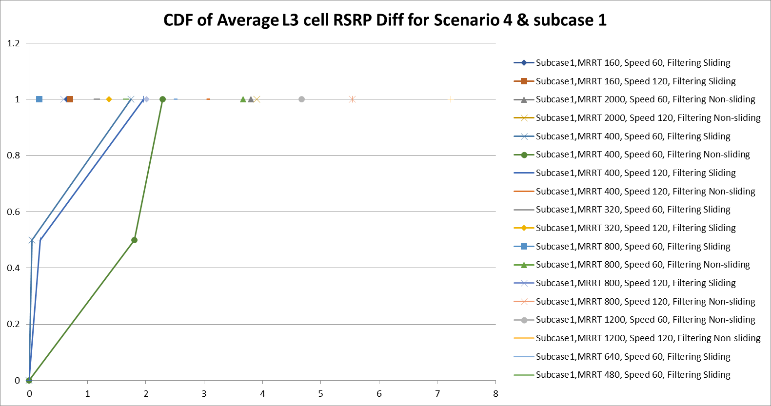
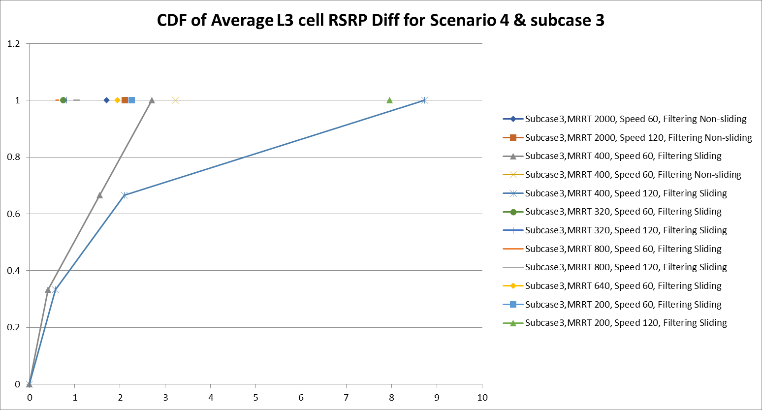
# Discussion

## Merge Sub-use Cases in the Graphic Illustration

Given that sub-use case is a significant setting in the evaluation of AI MOB, we initially attempted to plot separate CDF graphs for each sub-use case in temporal case B, inter-frequency, and temporal case A, as shown in Figure 1. However, we found that for sub-use cases 1 and 3, the limited quantity of results led to numerous group settings for which only one company has provided data. Consequently, this precludes the generation of CDF curves for these settings. Therefore, we have consolidated the results from all sub-cases to generate comprehensive CDF curves that represent the overall outcomes.

**Figure 1(a). CDF of Optimal Average L3 cell RSRP difference for sub-use case 1 and 3 in Temporal case B**

**Figure 1(b). CDF of Optimal Average L3 cell RSRP difference for sub-use case 1 and 3 in Temporal case A**

**Figure 1 Separate CDF curves for sub-use cases**

#### **Q1: Companies are invited to provide feedback on the consolidation of results from all sub-cases to generate the CDF curves, as opposed to generating CDF curves for each sub-case individually. Please share any concerns or suggestions you may have regarding this approach.**

|  |  |
| --- | --- |
| Company | Comment |
| Ericsson | Propose to use CDF for Sub-case 2 and use bar chart for the other sub-cases (1 and 3).  The CDF figures could have a finer granularity showing the results from more companies (in some figures, e.g. for event prediction, it looks like only the results from two ~ three companies are shown).  MediaTek: We can observe that for many settings with subcase 1 and 3, only 1 or few companies provide valid results, and it’s unable to display results trends and get any conclusion from the sparse results (use bar chart way). |
| ZTE | Agree |
| OPPO | Agree |
|  |  |
|  |  |

## Temporal Case B (Scenario 2)

RAN2 has agreed on the key parameters and settings for temporal case B, which include MRRT and UE speed. The UE speed is set to 30 km/h and 90 km/h, while MRRT is set to 50%, 66%, 80%, and 90%. Additionally, the filtering options with sliding and non-sliding are listed separately.

We have illustrated the CDF of key parameters and settings into two figures, as shown in Figure 2(a) and 2(b). Please note that only one company has provided results for MRRT=90%, making it difficult to discern a trend from the graph based on this single data point. Therefore, MRRT is set to 50%, 66%, and 80% in Figure 2 for graphic illustration. Since some companies have provided results for 60 km/h, to reflect the contributions of all companies, 60 km/h should also be considered when plotting the CDF curves.

**Figure 2(a). CDF of Optimal Average L3 cell RSRP difference for temporal case B with Sliding filtering**

**Figure 2(b). CDF of Optimal Average L3 cell RSRP difference for temporal case B with Non-sliding filtering**

**Figure 2 CDF of Optimal Average L3 cell RSRP difference for temporal case B**

#### **Q2: Companies are invited to provide feedback on the graphic illustration for temporal case B, specifically Figure 2. Please share any comments or suggestions you may have.**

|  |  |
| --- | --- |
| Company | Comment |
| Ericsson | Propose to divide the data based on speed and have one figure per speed. E.g. one figure for 30Km/h another for 60Km/h and another for 90Km/h  MediaTek: We can consider having one figure per speed when we evaluate other key parameters (e.g., MRRT or PW). The above chart we showed is just an overall chart, we will also plot the CDF curves of key parameters such as speed or MRRT or PW separately. You can find these figures in the summary spreadsheet. |
| ZTE | In the previous meeting, we have agreement that:  Average L3 cell RSRP difference and last predicted point L3 cell RSRP difference of measurement results within PW is captured in TR.  So, besides the figure for average RSRP difference, some figures for last predicted result within prediction window shall also be included. And we think their data quantity is enough to plot a figure, e.g. in MRRT 50%, Speed = 30/60/90km/h, Non-sliding filtering option.  MediaTek: Added in summarized spreadsheet for scenario 2. Technically, we can plot the data for the last point. We have not done so only because many companies did not submit or **directly** provide these results in “last predicted point” column. |
| OPPO | From the current Figure, it looks like UE speed has more impact compared to MRRT. In legend, if UE speed is listed as first dimension, then mixing UE speeds could help better understand. No strong opinion on ZTE’s comment. I believe the trend will likely the same i.e. there is not so much additional information.  MediaTek: We also plot the figure for mixing UE speed with fixed MRRT, this help to see the UE speed impact for prediction accuracy. For example: |
|  |  |
|  |  |

## Inter-frequency (Scenario 3)

For Inter-frequency, RAN2 agrees to use the cluster setting for graphic illustration. We plotted a comparison between clusters and cells extracted from the results of all companies in Figure 3.

**Figure 3. CDF of Optimal Average L3 cell RSRP difference for cluster vs cell**

#### **Q3: Companies are invited to provide feedback on the graphic illustration for inter-frequency prediction, specifically Figure 3. Please share any comments or suggestions you may have.**

|  |  |
| --- | --- |
| Company | Comment |
| Ericsson | It looks fine. |
| ZTE | agree |
| OPPO | Agree |
|  |  |
|  |  |

## Temporal Case A (Scenario 4)

RAN2 has agreed that for the graphic illustration of temporal case A, the key parameters are [OW:PW] or PW, and UE speed. We attempted both illustrations with [OW:PW] and PW, as shown in Figures 4 and 5.

**Figure 4(a) [OW:PW] with Sliding**

**Figure 4(b) [OW:PW] with Non-sliding**

**Figure 4. CDF of Optimal Average L3 cell RSRP difference for temporal case A with [OW:PW]**

**Figure 5(a) PW with Sliding**

**Figure 5(b) PW with Non-sliding**

**Figure 5. CDF of Optimal Average L3 cell RSRP difference for temporal case A with PW**

Upon comparing the two plotting methods, OW:PW and PW, we found that in the OW:PW method, the value of OW:PW does not correspond to a unique pair of OW and PW values. For example, both OW=PW=40ms and OW=PW=1600ms would be counted as OW:PW=1. This leads to an exceptionally wide range of results for OW:PW=1 in the final analysis, resulting in poor comparability and making it difficult to observe the impact trend of OW:PW on the results. Therefore, we suggest considering the impact of the key parameter PW on the results when plotting for temporal case A.

#### **Q4: Companies are invited to provide feedback on the graphic illustration for temporal case A, specifically Figure 5. Please share any comments or suggestions you may have.**

|  |  |
| --- | --- |
| Company | Comment |
| Ericsson | Propose to divide the data based on speed and have one figure per speed to make the figures more readable.  MediaTek: same as the response Q2. |
| ZTE | We agree to consider PW length as key parameter.  And, as our comment in Q2, the figure for last predicted point shall also be included.  MediaTek: same as the response Q2. |
| OPPO | It seems UE speed has more impact than PW. Maybe we can improve legends i.e. to order UE speed first. Plus for non-sliding option, it doesn’t make sense to split among UE speed. And no strong opinion on last predicted point.  MediaTek: For Non-sliding, we can update the figure as following: |
|  |  |
|  |  |

## Generalization

The generalization study cases with Configuration #A and Configuration #B including studies over UE speeds, and two sets of parameters (ISD, BS antenna height, BS Tx power) for cell configuration, represented by Configuration#1 and Configuration#2. For generalization, there is no need to duplicate all CDF curves of all key parameter settings for baseline, CG1 and CG2. We can fix one set of key parameters to illustrate the performance of baseline, CG1 and CG2 for UE speed and cell configuration. The fix set of key parameters is the one which most companies provide simulation results.

### **Generalization for Temporal Case B (Scenario 2)**

We choose the most common setting provided by companies that MRRT is set to 50%. Generalization on different UE speed (30/60/90 km/h). We can observe that, from Figure 6(a), the CDF curves are consistent at different speeds, which indicates that AI generalizes well at different speeds and that is consistent with the current agreement. For Figure 6(b), when MRRT is set to 50%, baseline UE speed=30km/h, Configuration #A is Configuration#1 and Configuration #B is Configuration#2, generalization on different cell configuration. We can observe that, the curves are consistent at different cell configurations, which indicates that AI generalizes well at different cell configurations and that is consistent with the current agreement.

**Figure 6(a) Generalization on UE speed**

**Figure 6(b) Generalization on cell configuration**

**Figure 6 CDF of Optimal Average L3 cell RSRP difference for temporal case B generalization performance**

#### **Q5: Companies are invited to provide feedback on the graphic illustration for the generalization of temporal case B (Scenario 2), specifically Figure 6. Please share any comments or suggestions you may have.**

|  |  |
| --- | --- |
| Company | Comment |
| ZTE | Some simulation results may be missing, since in the last meeting, many companies provide the simulation results for cell configuration in case B.  MediaTek: There are 7 companies provide the simulation results for cell configuration in case B, here we only show one example group of settings in Figure 6(b), i.e., MRRT is set to 50%, baseline UE speed=30km/h, Configuration #A is Configuration#1 and Configuration #B is Configuration#2, and filtering option with non-sliding, for the generalization on different cell configuration. That is the reason the results showed in Figure 6(b) is limited. |
|  |  |
|  |  |
|  |  |
|  |  |

### **Generalization for Inter-frequency (Scenario 3)**

Inter-frequency generalization graphic illustration with Configuration #A is “4GHz=>2GHz ”, and Configuration #B is “2GHz=>4GHz ”, we plot the CDF curves of baseline, GC1 and GC2. We can observe that generalization using GC#2 always outperform that of GC#1 and the GC#1 case without any preprocessing based on the information of predicted frequency suffers from significant performance loss, which is consistent with the current agreement.

**Figure 7 CDF of Optimal Average L3 cell RSRP difference for Inter-frequency generalization performance**

#### **Q6: Companies are invited to provide feedback on the graphic illustration for the generalization of inter-frequency (Scenario 3), specifically Figure 7. Please share any comments or suggestions you may have.**

|  |  |
| --- | --- |
| Company | Comment |
| ZTE | In the figure 7, why the number of companies providing the data for GC1 is larger than baseline. Companies are always expected to provide the corresponding baseline results.  MediaTek: That’s because some company only provide GC1 without baseline. (Can check the summarized simulation table). |
|  |  |
|  |  |
|  |  |
|  |  |

### **Generalization for Temporal Case A (Scenario 4)**

When PW is set to 400ms, we can observe that, the curves are consistent at different speeds, which indicates that AI generalizes well at different speeds, which is consistent with the current agreement. In the context of evaluating generalization performance across different cell configurations, the analysis is currently constrained by data from only two companies. This limitation affects the ability to visualize the performance metrics.

**Figure 8 CDF of Optimal Average L3 cell RSRP difference for Temporal case A generalization performance (PW=400ms)**

#### **Q7: Companies are invited to provide feedback on the graphic illustration for the generalization of temporal case A (Scenario 4), specifically Figure 8. Please share any comments or suggestions you may have.**

|  |  |
| --- | --- |
| Company | Comment |
| ZTE | For generalization over UE speed on temporal case A in FR2, some simulation results maybe missing. Many companies provide the simulation results for the PW = 400ms.  MediaTek: The figure is updated. Thanks for the correction! |
|  |  |
|  |  |
|  |  |
|  |  |

## Measurement Event Prediction

### **Temporal Case B (Indirect)**

The UE speed is set to 30 km/h and 90 km/h, while MRRT is set to 50%, 66%, 80%. Graphic illustrations on the performance metrics of F1 score is provided.

**Figure 9 CDF of F1 score for indirect event prediction temporal case B**

Only two companies have provided comparisons between the baseline and HO modeling option3, which limits the data available for analysis. As a result, creating graphical illustrations of system performance can’t provide meaningful insight.

#### **Q8: Companies are invited to provide feedback on the graphic illustration of the F1 score for indirect event prediction in temporal case B, specifically Figure 9. Please share any comments or suggestions you may have.**

|  |  |
| --- | --- |
| Company | Comment |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

### **Temporal Case A (Indirect)**

Considering intermediate KPIs, we choose the most common setting that PW=320ms with sliding filtering the results include UE speed=60km/h, 90km/h and 120km/h. Graphic illustration on F1 score is as following:

**Figure 10 CDF of F1 score for indirect event prediction temporal case A**

#### **Q9: Companies are invited to provide feedback on the graphic illustration of the F1 score for indirect event prediction in temporal case A, specifically Figure 10. Please share any comments or suggestions you may have.**

|  |  |
| --- | --- |
| Company | Comment |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

For system performance, we choose the most common setting that PW=320ms, UE speed=90km/h with sliding filtering option, we can observe that for Option 1 and Option 2 outperform legacy solutions in terms of HO failure number per UE per second and HOF rate, which is align with our current observation.

**Figure 11(a) CDF of HOF rate for Indirect event prediction temporal case A**

**Figure 11(a) CDF of HOF number per UE per second for Indirect event prediction temporal case A**

**Figure 11 System level performance for indirect event prediction temporal case A**

#### **Q10: Companies are invited to provide feedback on the graphic illustration of system performance for indirect event prediction in case A, specifically Figure 11. Please share any comments or suggestions you may have.**

|  |  |
| --- | --- |
| Company | Comment |
| OPPO | For Figure 11(a), maybe it makes more sense to show the difference instead of HF rate directly.  MediaTek: In current agreements on measurement event prediction, we have the comparison between option1/option2 and baseline, also has the comparison between option1 and option2. If we show the difference, it is difficult to visually display these two sets of differences at same figure. The curve that directly show the HOF rate is the easiest way.  1. Option 2 and Option3 outperform legacy solutions in terms of HO failure number per UE per second  2. Capture that most companies show in their simulation results that option 3 outperforms option 2 when prediction accuracy is good enough. There are some companies that show the opposite and indicate why. |
|  |  |
|  |  |
|  |  |
|  |  |

**Summary**:

Based on the limited feedback received from companies regarding the figures, the rapporteur assumes that the graphic illustrations are acceptable. Therefore, we will include these figures in the corresponding spreadsheets and propose to endorse the spreadsheets.

We will select the curves/settings to be included in the TR as needed.

# Conclusion

**Proposal 1: Endorse the performance evaluation spreadsheets as attachment to the TR.**

**Proposal 2: Endorse the CDF figures plotted based on the collected simulation results from companies in each individual spreadsheet.**

**Proposal 3: Select the CDF curves that align with the current observation trends to be included in the TR. The rapporteurs will choose the appropriate CDF curves as needed during the drafting of the TR.**

# Reference

### [1] R2\_129b\_ChairNotes\_25-04-11\_16-45 final