3GPP TSG RAN Meeting #101

RP-231965

AI/ML for Air Interface in Release 19

Bangalore, India, September 11-15, 2023

Agenda Item: 8A.2.1

Source: Nokia, Nokia Shanghai Bell

© 2023 Nokia

AI/ML for Air Interface in Release 19

Proposals for overall directions

1

Finalize Rel-18 study and TR for all the involved RAN WGs. Initiate Rel-19 work item based on the completed study and TR conclusions and recommendations

2

Study robustness of AI/ML enhancements in dynamic cellular radio environments for mobility use case(s) in Release 19

3

Study extensions to the AI/ML framework to enable reinforcement learning based AI/ML enhancements for air-interface in selected use case(s) in Release 19

© 2023 Nokia

Nokia’s views on the scope of normative work on AI/ML on air interface in Rel-19

© 2023 Nokia

AI/ML framework for air interface (1/2)

Enablers for ML air-interface derived from the studied Release 18 use cases

Specify support for solutions enabled by one-sided and two-sided models of proprietary format, including:

- Life-cycle management procedures
  - Air-interface signaling for Functionality based LCM, including functionality management, (de)activation/switching/fallback, and performance monitoring.
  - Air-interface signaling for a 'light' model-ID based LCM performed at the network side only for selected sub use cases, including
    - Identification of UE ML model(s) using model ID and metadata (if/when needed)
    - UE ML model (de)activation/switching/fallback (if/when needed)
  - Data collection procedures for training, inference and monitoring (with use case specific details as identified in Rel 18 SI)

- UE ML-related capabilities signaling mechanisms
  - Required enhancements of UE-capability reporting framework to enable functionality identification
  - Handling of UE operations that impact ML Functionalities and related LCM procedures

© 2023 Nokia

AI/ML framework for air interface (1/2)

Enablers for ML air-interface derived from the studied Release 18 use cases

- Mechanism to ensure continuity of ML-enabled functionalities during UE mobility
  - Handling of ML features/functionalities before, during and after UE mobility (cell handover or beam mobility), including data collection, ML inference, functionality monitoring, etc.
  - Required additional signalling for intra-gNB and inter-gNB mobility scenarios

- Robust performance and predictable UE behavior through UE requirements and test cases for UE Features enabled by one-sided and two-sided models
  - Pre/post deployment validation procedures
  - Generalization requirements as part of the performance requirements
  - Functionality-based LCM procedure requirements as part of core requirements
  - UE processing capability vs. number of ML-enabled features

© 2023 Nokia

Use case – CSI feedback enhancement

- Specify compressed CSI reporting based on a two-sided ML model, considering the following:
  - Assume offline model training with Network-first Type 3 separate training
  - Define mechanisms to ensure compatibility between the UE and NW-sided parts of the two-sided model
  - Ensure performance monitoring at the NW-side with possible assistance from the UE (if applicable)
  - Actions based on performance monitoring are limited to fallback to legacy codebook CSI feedback
  - Outline data collection procedures based on the existing Rel-16 Type II codebook CSI feedback
  - Enhance/reuse CSI reporting framework to support compressed CSI feedback
- Specify predicted CSI reporting based on a one-sided ML model at the UE, considering the following:
  - Reuse/enhance Rel-18 MIMO evolution CSI prediction framework to support ML-predicted CSI feedback
  - Ensure performance monitoring at the NW-side with possible assistance from the UE (if applicable) and incorporate the reuse of legacy CSI feedback
  - Ensure accurate data collection procedures at the UE side

© 2023 Nokia

Use case – Beam prediction

Specify beam prediction sub-use cases (spatial domain beam prediction and temporal domain beam prediction) with the UE ML model(s), considering the following:

- Support DL Tx or DL Tx-Rx beam prediction for a UE-sided model.
- For UE-sided beam prediction,
    - Define functionality LCM based on the beam measurement and reporting framework to support reporting of predicted beams.
    - Ensure mechanisms to select compatible model at the UE to support beam prediction, using functionality framework or model identification/LCM framework.
    - Enhance TCI state framework for the temporal domain beam prediction to enable beam indication for multiple time instances in future. Reuse legacy TCI state framework for spatial domain beam prediction.
- Ensure performance monitoring at the NW-side with possible assistance from the UE (if applicable) and incorporate the reuse of legacy beam measurement and reporting framework.
- Outline data collection procedures for based on legacy/enhanced beam measurement and reporting framework for UE.

© 2023 Nokia

Use case – Positioning enhancement

Specify AIML positioning enhancement sub-use cases at the UE and LMF one-sided ML model(s), considering the following:

- Prioritizing UE-side model for Direct and Assisted AI/ML positioning
- Functionality-based LCM to support AIML positioning at least for UE-sided model
- Assistance signaling (including report/request) and procedure to facilitate LCM operations:
    - Assistance signaling and procedure to facilitate generating and collection of training data such as Ground Truth (GT) label, model inputs, associated information of GT (quality indicator), measurement corresponding to model inputs etc.
    - Performance monitoring at the LMF-side with possible assistance from the PRU/UE, gNB (if applicable).
- Reuse legacy positioning and reporting framework as a baseline to support AI/ML positioning for both UE- and LMF-side model
    - Outline data collection procedures based on legacy/enhanced positioning measurement and reporting framework for PRU/UE and LMF.

© 2023 Nokia

New AI/ML studies in Rel.19

© 2023 Nokia

Further AI/ML studies for Air Interface

Motivation for AI/ML Framework Extension to Inter-Cell Mobility

Potential Benefits of ML-optimized mobility:

- Rel-18 mobility methods are well-optimized and in general perform well
- AI/ML potential in
  - Improving the inter-cell mobility performance in problematic scenarios
  - Reducing the amount of radio measurements and reporting that are required and optimizing resource usage
- ML can learn the mobility context, which enables tailored mobility solutions, for example for each cell boundary and/or UE type (fast or slow moving, trajectory etc.), or even for each individual handover

Challenges:

- Generalized vs. Localized ML models, i.e., models that are trained with training data from a limited scope or context
  - Potentially high number of model instances to be created and maintained
  - Model switching may be required
- Robustness and scalability of ML solutions

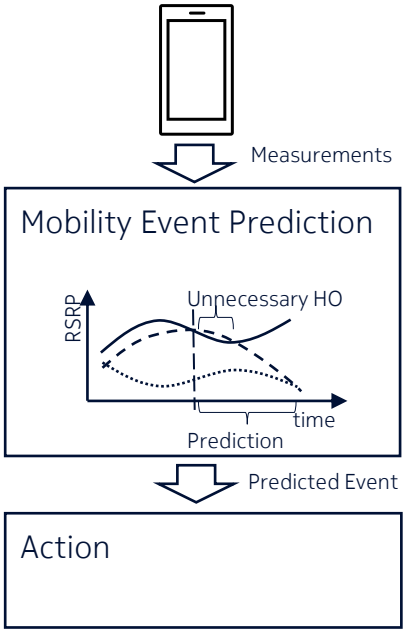
Build on top of Rel18 AI/ML framework and use cases

 © 2023 Nokia

Further AI/ML studies for Air Interface

Objectives for Rel19 Study on ML for Mobility

- Study inter-cell mobility optimization based on machine learning
  - Extending the ML concepts from Release-18
- Mobility event and signal prediction
  - How accurately selected mobility events or future UE measurements can be predicted?
  - Evaluate different optimization actions to be taken based on the predictions and their impact on network performance metrics.
  - Identify the most promising use case options.
  - Study the deployment options of the selected use cases, including
    - Scalability and transferability to different network deployments
    - Generalized or localized models
    - The robustness of the ML based mobility solutions.
  - How to evaluate and compare the solutions, e.g., the simulation scenario to be used
- Identify key requirements
  - Extensions to the Rel18 AI/ML framework
    - What are the deployment options and the required collaboration modes?
  - Any other RAN standardization impacts to support the use case(s)



Build on top of Rel18 AI/ML framework and use cases

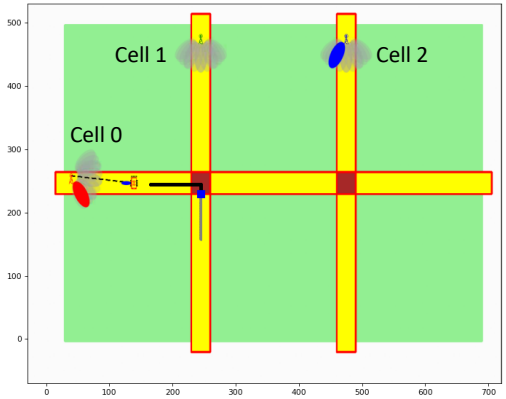
© 2023 Nokia

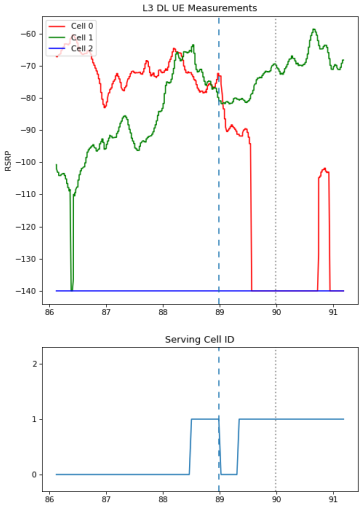
Further AI/ML studies for Air Interface

Study objectives for ML for Mobility

Use-cases

- Ping-pong and short-stay handover prediction to:
  - Reduce number of unnecessary handovers
  - Reduce the cost of ping-pong or short stay handovers
- For example, a street canyon situation may lead to either necessary or unnecessary ping-pongs depending on the UE type, trajectory and speed
- Machine learning can learn to predict such events for each UE based on the UE radio measurements and other inputs
- The models will need to learn the local radio environment and mobility patterns/trajectories, which is likely to require localized models





Build on top of Rel18 AI/ML framework and use cases

© 2023 Nokia

Further AI/ML studies for Air Interface

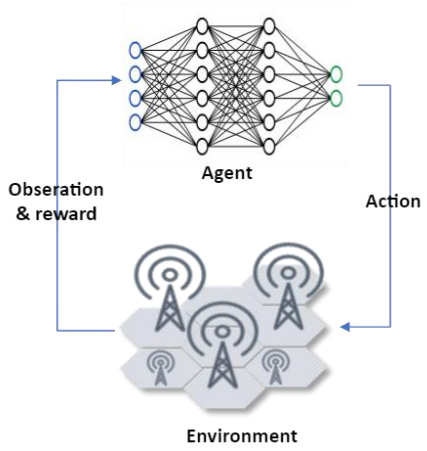
Motivation to AI/ML framework extension for enabling RL based enhancements

Benefits of (deep) Reinforcement Learning (RL) studies:

- Future-proof ML general framework, ready to evolve towards 6G.
- Enables goal-based online decision-making use-cases with arbitrary performance objectives (e.g., throughput or latency).
- RRM functions can be finetuned and adapted to the local environment with unique tailored solutions.

Key requirements for Reinforcement Learning in air-interface:

- **Exploration**: enabling exploration of the action space to train the RL model while minimizing performance degradation and overhead
- **Data collection**: access to training, inference and monitoring data required by the online RL models.
- **RL functionality/model-based LCM**: identify and manage different phases of an online RL model as it evolves with training



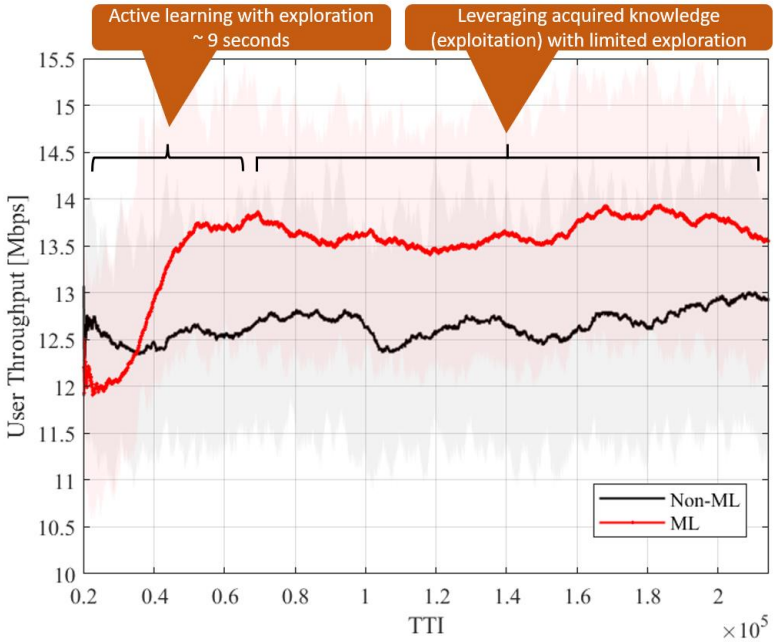
RL models are trained based on its interactions with the environment. Hence, training and data collection are performed while the RL model is deployed in air interface (online). This also cause the RL model to change its behavior while deployed in air interface. E.g., from exploration emphasized initial training to exploitation emphasized trained model.

Ensure future-proof ML framework towards 6G

 © 2023 Nokia

Further AI/ML studies for Air Interface

Motivation to AI/ML framework extension for enabling RL based enhancements



Example - beam management

- RL based approach to optimize DL Tx beam selection for user throughput.
- The goal of RL agent is to maximize the reward which is calculated based on user throughput.
- RL agent learns a better DL Tx beam selection policy tailored for a specific network configuration and environment to improve the user throughput.
- Improved user throughput around 10% after the active learning phase.
- Impacts user throughput during the active learning phase of RL agent due to exploration.

Unlock the potential of RL with future-proof AI/ML framework

© 2023 Nokia

Further AI/ML studies for Air Interface

Study objectives for Reinforcement Learning

Use-cases

- Prioritize beam management (UE-sided or NW-sided) as extension of Rel18
  - Other new use-cases are not precluded (e.g., scheduling, uplink power control, link adaptation)
- With selected use-cases
  - Study the performance benefits of RL
  - Study feasibility of UE and NW sided models
  - Identify AI/ML evaluation methodology for RL
  - Assess specification impacts to AI/ML framework (including phy/protocol (RAN1/2), interoperability and testability (RAN4)) as well as other case specific specification impacts.

General aspects of RL

- Identify key requirements and extensions to Rel18 AI/ML framework to enable RL based functionalities. Rel18 AI/ML framework is considered as the starting point.
  - Identify common notation and terminology.
  - Identifying key differences of RL functionality LCM compared to Rel18 AI/ML LCM.
    - Identify and define different phases (initial training, trained model, retraining) for RL functionality management
    - LCM operations specific to RL functionalities (if any).
  - Data collection for inference, training and monitoring of RL based functionality.
    - Enabling exploration of the action space with minimal performance degradation.
    - Training data collection for online RL models.

Study feasibility of Rel18 AI/ML framework to enable RL use-cases

 © 2023 Nokia

NOKIA