**3GPP TSG-SA5 Meeting #154S5-242036**

**15 - 19 April 2024, Changsha, Hunan, China**

**Source: Nokia, Nokia Shanghai Bell**

**Title:** **pCR 28.866 MDA management data correlation analytics**

**Document for: Approval**

**Agenda Item: 6.19.2**

# 1 Decision/action requested

***The group is asked to discuss and agree on the proposal.***

# 2 References

[1] 3GPP TR 28.866-000 “Study on Management Data Analytics (MDA) – Phase 3”.

# 3 Rationale

The data from different aspects of the network are correlated, that in many cases real analytics value comes from leveraging the correlation among the data. This pCR is to add an MDA use cases on correlation of data for different uses and from different sources.

# 4 Detailed proposal

|  |
| --- |
| **1st Change** |

# Y Use cases and requirements for MDA capabilities and services

# Y.2.x Data correlation analytics

#### Y.2.x.1.1 Description

This MDA capability is for data correlation analytics on network, computing and slice resources across different aspects consisting of network domains(RAN/ core/ transport), resource domain (network and cloud) and slice / slice subnet aspects.

#### Y.2.x.1.2 Use cases

##### Y.2.x.1.2.a Measurement data correlation analytics for ML training

For a ML model training, a large amount of measurement data instances may be collected and does not necessarily add value to training performance. The collected measurement data for ML training can be highly correlated (linear or non-linear), hence there may be high redundancy among the collected measurement data. Hence using all measurement data for a model training (and inference) can be a waste of computing resources and energy. Therefore, it’s necessary to optimise the training data preparation based on the correlation analytics and redundancy information. A correlation analytics may help to detect the redundancy pattern among the measurement data for ML training, such as:

- For a given task (e.g., an analytics, model training), analyses the correlation among the given set of measurement data, the output can be a (much) smaller set of measurement data, with which ML model (re-)training could be much more efficient with limited (or managed) impact to model training performance (comparing to use full set of data). The output may include a recommendation, for example recommendation to optimize measurement data collection for the model training.

- Regularly renew the correlation analytics as time progresses, since the correlation relationship might change; this is especially useful when there is a need to regularly re-train the ML model.

#### Y.2.x.1.3 Potential Requirements

Table Y.2.x.1.3-1

|  |  |  |
| --- | --- | --- |
| Requirement label | Description | Related use case(s) |
| **REQ-MDA\_CORE-01** | MDA capability for Measurement data correlation analytics for ML training should include the capability to provide the measurement data redundancy analysis including which measurement data correlate to which measurement data, the rate of redundancy, and recommendation to optimize measurement data collection for the model training. | correlation analytics (clause Y.2.x.1.2.a) |

#### Y.2.x.1.4 Potential Solutions

##### Y.2.x.1.4.1 Possible solution for measurement data correlation analytics for ML training

* Introduce a data type for measurement data correlation analytics recommendation, called measurementDataCorrelationRecommendation. The data taype can be the contents of the analytics report representing the recommendations from MDA for the measurement data correlation analytics for ML training.
	+ The contents of this data type may be a set of 3GPP and non-3GPP recommendations.
* An optional attribute may configure the analytics context in MDARequest. The context may include attributes
	+ An optional attribute may indicate the training performance requirement.
	+ An attribute may configure the scenario type of correlation as an enumeration.

|  |
| --- |
| **End Change** |