**3GPP TSG-SA5 Meeting #162 *S5-253881d1***

Goteborg, Sweden, 25 - 29 August 2025

**Source: China Mobile**

**Title: Rel-19 pCR TS 28.561 Update NDT data generation**

**Document for: Approval**

**Agenda item: 6.19.5.1**

**Spec: 3GPP TS 28.561**

**Version: 1.0.0**

**Work Item: NDT**

**Comments**

This contribution is proposed to update NDT data generation to improve quality.

**Proposed Changes**

\* \* \* First Change \* \* \* \*

## 5.4 NDT support for data generation

### 5.4.1 Description

NDTs can be used to support many generation capabilities for different application use cases.

### 5.4.2 Use cases

#### 5.4.2.1 General use case on NDT support for data generation

The NDT should support a capability to provide a report/output on the simulation/emulation enabling generation of data and information related to network scenarios, configurations, policies, and performance outcomes. When receiving the request from the MnS consumer, the MnS Producer creates the NDT job and determines the network objects and data (e.g., PM data, CM data) according to data requirements. The simulated network objects include simulated network areas and managed objects in NDT. And then, according to the modelled network objects and data, the MnS Producer collects the data from the physical network used for NDT job simulation. Subsequently, based on the collected data, the MnS Producer creates/activates NDT instance, executes the network simulation and generates data corresponding to the request by using the NDT.

The data generation scenarios that the NDT might support include those in following sub-clauses.

#### 5.4.2.2 Using NDT to generate ML training data

ML training usually requires large amounts of data to guarantee good performance of the ML models. In general, the ML training data for network related use cases is obtained through historical network management data. For instance, assuming that there is a ML model supporting MDA SLS analysis described in TS 28.104 clause 7.2.2, the raw feature of training data could be the enabling data, such as UL/DL throughput, uplink/downlink delay, etc., as specified in clause 8.4.2 of TS 28.104.

However, obtaining data from the network has the following limitations:

- The quantity of issues happened in actual mobile network is limited.

- The variety of issues happened in actual mobile network is limited. There could be corner network issues that hardly happen in live network.

Sufficient ML training data plays a key role to a useful ML model. The more training data provided, the better the performance of ML model. To overcome these challenges, the MnS consumer can request the NDT to generate data with an indication of data requirements, e.g. data type, required data period, data sampling periods, etc. After the NDT modelling, the MnS producer sends a report with the generated data to the MnS consumer, which can be used to enhance model accuracy by providing a wide range of training examples reflecting potential real network conditions.

\* \* \* End of Changes \* \* \* \*