**3GPP TSG SA WG 1 Meeting #111 S1-253231r2**

**25-29 August 2025, Goteborg, Sweden** *(revision of S1-25xxxx)*

**Source: Huawei, China Mobile, CATT, Nokia, ZTE, NTT DOCOMO**

**pCR Title: Update of clause 5.9.2 use case on efficient data collection and control for 6G system**

**Draft Spec: 3GPP TR 22.870 v****0.3.1**

**Agenda item: 8.1.2**

**Document for: Approval**

**Contact: Fangyuan ZHU, zhufangyuan@huawei.com**

*Abstract: this contribution proposes to update the existing clause 5.9.2 use case on efficient data collection and control for 6G system.*

**1. Introduction**

This contribution proposes to update the existing clause 5.9.2 use case on efficient data collection and control for 6G system

**2. Reason for Change**

This contribution merges the following contributions

|  |  |
| --- | --- |
| [S1-253118](file:///C:\TSGS1_111_Goteborg\Docs\S1-253118.zip) | China Mobile |
| [S1-253189](file:///C:\TSGS1_111_Goteborg\Docs\S1-253189.zip) | CATT |
| [S1-253206r1](file:///C:\TSGS1_111_Goteborg\docs\S1-253206r1.zip) | Nokia |
| [S1-253231r1](file:///C:\TSGS1_111_Goteborg\docs\S1-253231r1.zip) | Huawei |
| [S1-253246](file:///C:\TSGS1_111_Goteborg\Docs\S1-253246.zip) | ZTE Corporation, China Unicom |
| [S1-253253](file:///C:\TSGS1_111_Goteborg\Docs\S1-253253.zip) | NTT DOCOMO |

**3. Proposal**

It is proposed to agree the following changes to 3GPP TR 22.870 v0.3.1.

\* \* \* Next Change \* \* \* \*

# 2 References

The following documents contain provisions which, through reference in this text, constitute provisions of the present document.

- References are either specific (identified by date of publication, edition number, version number, etc.) or non‑specific.

- For a specific reference, subsequent revisions do not apply.

- For a non-specific reference, the latest version applies. In the case of a reference to a 3GPP document (including a GSM document), a non-specific reference implicitly refers to the latest version of that document *in the same Release as the present document*.

Editor's Note: all References numbers to be corrected, missing references to be added

[r1] <https://portal.etsi.org/webapp/WorkProgram/Report_WorkItem.asp?WKI_ID=74875>

\* \* \* Next Change \* \* \* \*

### 5.9.2 Efficient data collection and consumption for 6G system

#### 5.9.2.1 Description

To better serve the users and manage the network, as well as provide non-connectivity service (e.g. LCS), 3GPP system need to conduct collection of data associated with the network and services provided by the 3GPP system. The data collection in the 5GA network for network/service management and non-connectivity service is defined for specific use cases, resulting in varying solutions across different use cases, different layers and different domains. Here is a brief summary of different data collection work done in RAN/SA2/SA4 and corresponding security mechanism in SA3:

- In RAN domain, data collection from UE is based on MDT (Minimization of drive testing) method as defined in TS 37.320 [237] and TS 38.331 [238]. The user consent of UE location data collection for MDT is defined in TS 32.422 [239].

- For data collection for AI/ML in core network, the data collection feature permits NWDAF to retrieve data from various sources (e.g. NF such as AMF, SMF, PCF, NSACF, GMLC, and AF, OAM, etc.), as a basis of the computation of network analytics as defined in TS 23.288 [114]. The operator can leverage these network analytics to optimize the management of network operations. For instance, as outlined in clause 6.22 of TS 23.288 [114], the operator can collect network data (such as OAM data and CP signalling) to detect or predict a signalling storm and use the analytics to mitigate or prevent its impact. The data collection is performed via SBI between CN NFs, DCCF (Data Collection Coordination Function) is introduced as logical functionality to coordinate data collection in CN, and ADRF (Analytics Data Repository Function) is defined to store the collected data. The user consent of UE data collection for AI/ML is defined in TS 33.501 [250].

- For UE positioning, the input data is collected from UE/RAN to CN via CP/UP path as the basis of the computation of UE location as defined in TS 23.273 [240]. The privacy check of UE location collection is also defined in TS 23.273 [240].

- To collect data and to expose data between CN domain and AF domain, the network exposure mechanism in core network has been defined in clause 5.20, TS 23.501 [140] and clause 4.15, TS 23.502 [30], where the NEF is used as the termination node to isolate both sides. Especially for edge, exposure of UE data is defined in TS 33.558 [241].

- If the data collection is triggered by OTT server in application layer from UE, the UP path may be used with the assistance of DCAF, as defined in TS 26.531 [242].

The aforementioned mechanisms are defined for different network functions to support data collection in different scenarios, leading to high standardization overhead. There is lack of coordination between these network functions, leading to isolated data collection and duplicated data collection. Meanwhile, the data transmission based on the control plane are not suitable for transmitting large amounts of data, leading to inefficient data transmission. Therefore, the data collection and control mechanism need to be improved in 6G.

Table 5.9.2.1-1 lists several potential use cases for data collection in 6G, and illustrates the corresponding data type, data provider, data consumer and data volumes for each use case. Take AI/ML data as example, there is a large amount of data types defined in 3GPP apart from those non-standardized ones, and the scale of data volume is usually large (from 10k to 100M).

Table 5.9.2.1-1: Heterogeneous 6G System Data

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Use case category** | **Data types** | **Data provider** | **Data Consumer** | **Data volume** |
| Sensing data collection/sensing result exposure  (e.g. measurement) | Sensing measurement data requires further study and discussion. The potential sensing measurement can be channel estimation results, delay, Doppler, angle, signal strength, etc. | UE, RAN | Sensing Function | Related to use cases: e.g. data rate of measurement is about 10Mbps for objects creating hazards on roads |
| AI/ML data collection / analytics exposure  (e.g. AI/ML for air interface, AI/ML for core network) | - For RAN data collection for Air interface, see TR 38.843 [243]  - For core network data collection, see the data collection framework in clause 6.2, TS 23.288 [114], and the detailed data type are defined from clause 6.3 to clause 6.21, TS 23.288 [114]. | CN, RAN, UE, AF | CN, RAN, UE | Typically, the internal data collection for AI model training yields tens of thousands of samples.  Related to AI model: e.g. the data volume is about 10k bytes to 100M bytes for AI model of beam management |
| Energy Efficiency (EE information exposure to AF) | still under the discussion in SA2 R19 EnergySys WID | CN, OAM | EIF (Energy Information Function) | Related to use cases: e.g. could be similar to data volume for charging |
| Positioning, | See clause 6, TS 23.273 [240] | UE, RAN | LMF, UE | Related to use cases |
| Network Exposure in core network | See clause 5. 20, TS 23.501 [140], and clause 4.15, TS 23.502 [30] | 5GC NF, AF | 5GC NF, AF | Related to use cases |
| … |  | … | … | … |

All the above use cases for data collection and control share the following commonality：

1. The data collection and control mechanisms need to support procedures such as data collection, data storage, data exposure and data deletion.
2. The collected data could be requested or consumed by multiple data consumers within 3GPP system to support multiple services; for example, a single spatiotemporal environmental dataset can be concurrently utilized by multiple services, including IoT-enabled pervasive sensing service, latency-sensitive extended reality (XR) applications requiring <10ms motion-to-photon delays, and intelligent transportation systems (ITS) with centimetre-level geolocation precision.
3. The upper bond of the collected data volume is very high, e.g. on the order of 100 M byte;
4. Latency requirements for data collection and transmission exhibit variations across diverse use cases.
5. The collected data transmission path between different entities within 3GPP system can vary widely, encompassing scenarios such as UE to RAN node，UE to CN node, RAN node to CN node, and CN node to CN node;
6. All these use cases that involve collecting UE related data shall support security, privacy protection and user consent. In addition, security requirements shall consider different data types.
7. All these use cases that involve data exposure shall be subject to operator policy and user consent.

A screen shot of a computer

AI-generated content may be incorrect.

Figure 5.9.2.1-1: Use case specific data mechanisms in 5G 🡺 common data mechanism in 6G

As shown in Figure 5.9.2.1-1, use case specific mechanisms are defined for each service scenario in 5G, leading to system complexity and inefficient data collection. As 6G networks integrate advanced capabilities progressively, such as AI, enhanced computing power, multi-dimensional sensing and digital twin network, they will need to process massive amounts of multi-source, heterogeneous data (from diverse formats and data resource) generated by 6G system. Consequently, 6G system shall support efficient data collection and control mechanism capable of addressing both universal requirements and scenario-specific demands across heterogeneous use cases such as AI/ML, sensing, XR service and digital twin network. The efficient data collection and control mechanism shall support various data providers and consumers, which maintains service-agnostic characteristics while embedding extensibility for emerging service.

Also, for some cases that the collected data or generated data should be stored in 6G system with effectively management, in order to minimize redundant data collection. In order to fulfil the service performance requested by data consumer, how to guarantee the quality of data service should be considered.

In order to fulfil the service performance requested by data consumer, data quality should be considered across the entire data lifecycle [r1]. Different phases can utilize different metrics to characterize data quality. By considering those data quality metrics in different phases together, the E2E service performance will be guaranteed by 6G system. For example:

- Data generation phase: data freshness, data accuracy, and data reliability (e.g. accuracy error 0.1 m).

- Data collection phase: data volume (e.g. 500M byte).

- Data processing phase: data consistency (e.g. consistency 99%)

- Data transmission/consumption phase: transmission latency (e.g., 200ms).

In conclusion, the 6G system is expected to support the following characteristics:

- Common data framework that can serve the need of multiple use cases, to improve data collection efficiency and avoid duplication data collection

- Support data collection from various data providers and data exposure to various data consumers;

- Support user consent for data collection; and

- Support data exposure subject to operator policy and user consent.

#### 5.9.2.2 Potential New Requirements

Editor’s Note: user consent in this section is FFS.

[PR 5.9.2.2-1] The 6G system shall support mechanisms for 6G System Data collection and consumption minimizing the impact to 6G services.

[PR 5.9.2.2-2] Subject to operator’s policy, regulation and user consent, the 6G system shall support processing of 6G System Data.

NOTE: examples of data processing are use case dependant, e.g. data fusion, data anonymization and data analysis.

[PR 5.9.2.2-3] Subject to user consent, regulation and operator's policy, the 6G system shall support secure means to expose 6G System Data to authorized trusted third-party, authorized network function or authorized UE.

[PR 5.9.2.2-4] The 6G system shall support security and privacy protection of the 6G System Data.

[PR 5.9.2.2-5] The 6G system shall be able to provide charging and accounting mechanisms for 6G System Data.

[PR 5.9.2.2-6] The 6G system shall support data storage and data retrieval of 6G System Data.

[PR 5.9.2.2-7] Subject to regulation and operator policy, the 6G system shall support efficient transmission/distribution of 6G System Data between different data providers and data consumers.

[PR 5.9.2.2-8] Subject to regulation and operator’s policy, the 6G system shall support mechanisms to guarantee data quality and service performance across the entire data lifecycle, including data generation, data collection, data processing and data transmission.

[PR 5.9.2.2-9] Subject to regulation and operator policy, the 6G system shall support standardized data formats for the detection of network failures.

\* \* \* End of Change \* \* \* \*