**3GPP TSG-SA3 Meeting #108-Bis-e *S3-222927-r6***

**e-meeting, 10 - 14 October 2022** *revision of S3-222570*

**Source: Interdigital, Apple, AT&T, CableLabs, Convida Wireless, Deutsche Telekom, Ericsson, Intel, JHU, Google, Lenovo, Nokia, NCSC, NTT DOCOMO, Oppo, Philips International B.V., Qualcomm Incorporated, US NIST, US NSA, Verizon, Xiaomi, ZTE**

**Title: PCR for KI #1: Privacy aspects of variable length user identifiers**

**Document for: Approval**

**Agenda Item: 5.4**

# 1 Decision/action requested

***This document proposes cleanup to key issue #1: Privacy aspects of variable length user identifiers. SA3 is kindly requested to approve this document.***

# 2 References

[1] 3GPP TR 33.870 v0.1.0

# 3 Rationale

This PCR addresses three editor’s notes in KI #1.

Main reasons:

Backward Compatibility: Since Rel-16, subscriptions (e.g., NPN subscriptions) have already been assigned NAI SUPIs of variable and unequal length that reflect the names that the end-users of those NPN subscriptions are using. If we assume that operators will now move to a scheme where usernames are all of equal length, then this would require users to be assigned new SUPIs when they upgrade devices. The 3GPP will study and ultimately propose a solution that will allow to preserve already-assigned meaningful username IEs of SUPI in the NAI format.

Flexibility and preserving operators’ choice: Variable length username allows MNOs to take advantage of the flexibility of the NAI SUPI format. In many cases, MNOs use SUPIs that are provided by the NPN owners and would prefer to have variable-length usernames that won’t leak private information, i.e., a 3GPP-developed solution instead of managing the "username" field.

Editor’s Notes in clauses 5.1.2 and 5.1.3 are redundant. 3GPP is contribution-driven and new threats and requirements can be added upon contributing with or without having such Editor’s Notes.

Additional reasoning and information to support this proposal are in the accompanying discussion paper in S3-222580.

# 4 Detailed proposal

\*\*\*\* START OF CHANGE 1 \*\*\*\*

## 5.1 Key issue #1: Privacy aspects of variable length user identifiers

### 5.1.1 Key issue details

Networks can decide to allow user identifiers with variable length, e.g. in case SUPI of type NAI. If an attacker can learn something about the length, this will reduce the size of the anonymity set.

The length can become visible to an attacker in case a length preserving encryption scheme is being used for identifier concealment.

3GPP authentication schemes referred to in TS33.501[xx] are: 5G-AKA and EAP-AKA', which are mandatory to support, as well as other key generating EAP methods, e.g., EAP-TLS and EAP-TTLS. All of these methods identify the subscriber using SUPI. As SUPI of type IMSI has a fixed length, this key issue is not applicable to SUPIs of type IMSI.

For NAI based SUPI types, the authentication method may leak the length of the SUPI even if identifier privacy mechanisms specified for the authentication methods are used.

These privacy mechanisms are:

- For 5G-AKA and EAP-AKA' the mechanisms are profile A, profile B, or proprietary SUCI calculation scheme.

- When some EAP based methods are used, e.g., EAP-TLS and EAP-TTLS, an anonymous SUCI can be used, and the actual SUPI is sent after an EAP secure channel is established, e.g. the TLS tunnel.

### 5.1.2 Security threats

An attacker on the air interface can identify and track subscribers with unusual lengths of the username field of variable-length SUPI in NAI format even if it is confidentiality protected (e.g., relatively short or long SUPIs).

Note: NAIs can be used for any EAP method.

If such an unusual length of the username field is unique to a single subscriber, an adversary might be able to uniquely attribute it to that subscriber.

If there is a group of subscribers with unusual lengths of username fields in their SUPIs, the attacker might be able to infer the membership of those subscribers in such a group.

When using EAP methods for authentication, it is not sufficient to protect the variable-length SUPI in NAI format against the above threat only in NAS messages (e.g., protecting the SUPI when sending it in a Registration Request). When such variable length SUPIs (i.e., username) are also used in EAP authentication methods (irrespective of whether the EAP authentication method is privacy preserving or not), an attacker may be able to identify and track subscribers at the EAP layer even if the user identifier is protected in the NAS layer against the above attack. This is because an attacker may be able to perform the above attack by identifying the confidentiality protected NAI within the EAP message that is sent over the air and then inferring the length of the NAI even if it is ciphered. The attacker can perform the same attack actively by sending an EAP Identity request to the UE. For example, in case of EAP-TLS or EAP-TTLS, even if the identifier is sent after TLS ciphering is turned on, the attacker may be able to infer the length of the EAP identifier of the UE by locating the ciphertext associated with the identifier. This attack is possible since TLS (both TLS 1.2 and TLS 1.3) leaves any padding to the application. Moreover, the EAP-TLS RFC does not specify any such padding (RFC 9190 recommends use of padding only for TLS record packets to hide the length of client certificates, c.f., section 5.8 of RFC 9190).

NOTE: The above threat of using EAP layer to infer the length of NAI is not applicable for 5G EAP-AKA’ specified in TS 33.501. In 5G EAP-AKA’, the UE always sends the same SUCI in the EAP layer. 5.1.3 Potential security requirements

The 5G system should protect against anonymity set reduction based on identifier length.

Note: the following conditions are necessary for proper evaluation of a solution

* the solution needs to indicate which authentication mechanisms it works with and whether that authentication mechanism preserves SUPI length.
* the solution needs to be evaluated as to whether it is backwards compatible with SUPIs in NAI format, which might already be deployed.

\*\*\*\* END OF CHANGE 1 \*\*\*\*