**3GPP TSG-SA3 Meeting #108e-AdHoc *S3-222820-r2***

**e-meeting, 10 - 14 October 2022**

**Source: Ericsson**

**Title: Padding-based solution to the leakage of the length of SUPI through SUCI**

**Document for:** **Approval**

**Agenda Item: 5.4**

# 1 Decision/action requested

***In this box give a very clear / short /concise statement of what is wanted.***

# 2 References

[1] 3GPP TS 23.003: "Numbering, addressing and identification".

[2] 3GPP TS 33.501: "Security architecture and procedures for 5G System".

# 3 Rationale

According to TS 23.003 [1], subscribers’ permanent identifiers, called SUPIs, are allowed to be in Network Access Identifier (NAI) format (username@realm), which can have variable lengths.

Using the Elliptic Curve Integrated Encryption Scheme (ECIES), a user device and the home network agree on a shared key by leveraging the public key of the home network. The user device uses the shared key in a symmetric encryption scheme (AES in counter mode) to encrypt SUPIs, into concealed identifiers, called SUCIs [2].

In the symmetric-key setup, security notions like real-or-random, left-or-right, or semantic security are defined in the context where plaintexts have the same lengths [3]. Though AES counter mode is secure according to these notions, direct use of it is not sufficient to serve an intended purpose of SUCIs -- indistinguishability of SUCIs. This is because SUPIs can have different lengths, and in counter mode, the length of the plaintext and the corresponding ciphertext is the same. Therefore, when two SUPIs have different lengths, their ciphertexts are distinguishable from each other, causing the reduction of the associated anonymity set.

# 4 Detailed proposal

\*\*\* 1st 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*.

[1] 3GPP TR 21.905: "Vocabulary for 3GPP Specifications".

[2] 3GPP TS 24:501: “Non-Access-Stratum (NAS) protocol for 5G System (5GS)”.

[XX] 3GPP TS 33.501: "Security architecture and procedures for 5G System".

[XY] IETF RFC 7542: "The Network Access Identifier".

\*\*\* 2nd CHANGE \*\*\*

## 6.A Solution #A: Concealing length of SUPIs in SUCIs by padding the SUPIs

### 6.A.1 Introduction

This is a solution to KI #1. It uses pre-encryption padding of SUPIs.

EN: This solution may need to be updated to align with the KI once the ENs in the KI is resolved.

### 6.A.2 Solution details

#### 6.A.2.1 Solution Basics

For SUPIs taking the form of a NAI, the subscription identifier part of the SUPI includes the "username" portion of the NAI as defined in NAI RFC 7542 [XY]. With an exception for the null scheme, the UE pads the username portion before encrypting it using ECIES.

NOTE 1: The null scheme does not provide SUPI privacy in the first place; therefore, padding will only increase bandwidth without improving any privacy.

NOTE 2: In the present document, SUPI in IMSI format is not padded because it is fixed length, and SUCI cannot be attributed to a particular SUPI based on length.

The necessary padding parameters, which are the home operator's choices, is stored in USIM.

#### 6.A.2.2 Padding parameters

Padding parameters comprise a list of pLen. Each pLen in the list indicates the number of octets in the padded username for a certain number of octets in the unpadded username.

The HN stores the whole pLen. The USIM stores only one element of the pLen that indicates the number of octets in the padded username for the unpadded username associated with the USIM.

Editor’s Note: The complexity of recalculating effective padding parameters when the length distribution of SUPIs changes is FFS.

#### 6.A.2.3 UE Side

Padding of SUPIs in NAI format is performed by the same component, either USIM or ME, that performs the calculation of SUCI in the following manner:

- If the number of octets in the unpadded username is indicated in the list of pLen, the username is padded to the corresponding pLen. In this case, the username is prepended with the necessary numbers of octet value 0x20.

- If the number of octets in the unpadded username is not indicated in the list of pLen, the username is not padded.

- The UE does not pad the SUPI in NAI format when using the null scheme.

#### 6.A.2.4 Home Network Side

With an exception for the null scheme, when the de-concealed SUPI is in NAI format and padded by the UE, the SIDF unpads the username portion. The SIDF removes 0x20 octet from the beginning of the username until a non-0x20 octet value is identified.

### 6.A.3 Evaluation

Editor's Note: The k-anonymity analysis in the context of an IMSI catcher is FFS

\*\*\* END OF CHANGES \*\*\*